Statistical methods for history matching

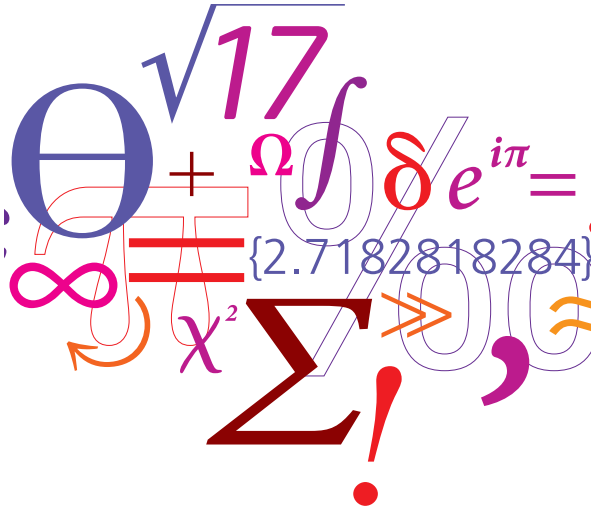
Johansen, Kent

Publication date:
2008

Document Version
Publisher's PDF, also known as Version of record

Technical University of Denmark

Statistical methods for history matching



Kent Johansen
2008

DTU Chemical Engineering
Department of Chemical and Biochemical Engineering

Statistical Methods for History Matching

Kent Johansen

Department of Chemical and Biochemical Engineering
Technical University of Denmark
Kongens Lyngby, Denmark

Technical University of Denmark
Department of Chemical and Biochemical Engineering
Building 229, DK-2800 Kongens Lyngby, Denmark
Phone +45 45252800, Fax +45 45882258
kt@kt.dtu.dk
www.kt.dtu.dk

ISBN 978-87-91435-75-7
Printed by J&R FRYDENBERG A/S, Copenhagen, Denmark

Acknowledgements

This work is submitted as a partial fulfillment of the requirements for obtaining the PhD degree at the Technical University of Denmark. The work has taken place at the Center for Phase Equilibria and Separation Processes (IVC-SEP) at the Department of Chemical Engineering, Technical University of Denmark (DTU). The work started in August 2004 and ended in October 2007. This period includes a two months leave where the BOSS reservoir simulator was extended with various features and a six months stay at Stanford University as a visiting researcher. The project is partly financed by DONG Energy, Forsker Uddannelsesrådet (FUR), and DTU in the MP_2T graduate school in chemical engineering framework.

I would like to thank PhD student Morten Rode Kristensen for being the perfect office mate. Our numerous discussions on numerical issues in reservoir simulations and other issues have been very valuable. Also, Morten's part in the development of the BOSS reservoir simulator should be acknowledged - the simulator would not have been so comprehensive without Morten's engagement. My thanks also goes to my supervisors Alexander A. Shapiro and Erling H. Stenby for being patient and understanding. I would like to express my gratitude to associate professor Jef Caers from Stanford University who made my six months stay at the Stanford Center for Reservoir Forecasting (SCRF) possible. My thanks goes to all the SCRF students and researchers who have assisted me during my stay. My special thanks goes to Satomi Suzuki for her enthusiastic and welcoming nature. I also thank all the Stanford people with whom I spent my free time. These persons made my stay memorable, not only with respect to my research, but especially with respect to all the adventures we made together.

The staff and students at the IVC-SEP are thanked for providing a welcoming and informal atmosphere.

Last, but not the least: Thanks to my dear Heidi for supporting me.

Kent Johansen, Kgs. Lyngby, Denmark. Oct. 2007

Summary

This dissertation represents three years work on the topic statistical methods for history matching of oil production. The work has resulted in a comprehensive treatment of existing techniques and the development of a novel technique which combines gradient information with a stochastic history matching method.

History matching is an important part of the management of an oil field and poses many problems due to the complexity of typical reservoir models and the shear size of the simulation model used to perform fluid flow simulations. The term *history matching* refers to the process of adjusting a reservoir model such that simulated production data agrees with actual measured production data. From a producing field there will be measurements of production data from the beginning of production until the present time. Such measurements over time are referred to as historic data. The reservoir model is adjusted until the production history is identical (or close) to the production calculated by a numerical fluid flow simulation. The reservoir model is mainly used to forecast future production which is crucial for many aspects of the field management and development. The reason for history matching the reservoir model is that it is thought to have better predictive capabilities if it at least describes the past production well. In other words: why would we trust a prediction from a reservoir model which does not even agree with historic production data?

The dissertation is divided into three parts which are subdivided into a number of chapters:

- Background and introduction
  - Introduction to history matching and literature review
  - Introduction to geostatistics
  - Introduction to reservoir simulation and adjoint sensitivity calculation

- Methods for history matching
  - Gradual deformation
  - Probability perturbation
  - Deterministic history matching
  - Hybrid method
- Appendix with details on various topics

In the first part a detailed introduction to geostatistics is supplied. Geostatistics makes up the backbone of many history matching methods, including the two stochastic methods used in the dissertation. A solid understanding of the geostatistical background of these methods is important. One other issue is important not only for geostatistically based algorithms but for all issues related to history matching: reservoir simulation. The mathematical background for reservoir fluid flow simulation is presented together with a more detailed treatment of the numerical solution of the discretized flow equations. Efficient calculation of sensitivities of production data with respect to grid block properties (e.g. permeability) is a central issue in the use of deterministic methods as well as in the proposed hybrid method. Therefore, a comprehensive treatment of the adjoint sensitivity calculation is provided in the chapter concerning reservoir simulation. Implementational details are discussed in further detail in the appendix.

The part of the dissertation related to history matching methods is divided into four chapters which deal with individual methods. The last chapter is devoted to the introduction and discussion of a novel technique which combines an existing geostatistical method with gradient information.

The methods which are discussed in the dissertation are exemplified in cases where they are applied to history matching problems. Emphasis is put on exemplifying the proposed hybrid method since it is a new approach and as such is not dealt with in previous works. The hybrid method has shown to improve the performance of traditional probability perturbation. The work has been limited to binary facies models. However, the method should be extendable to more complex cases where more facies are present. The proposed method seeks to improve the convergence of the probability perturbation method by including qualitative gradient information. The qualitative information is extracted from the gradient of the objective function by a simple filter which filters out numerically small elements of the gradient. A new parameter denoted *degree of trust* is introduced. This parameter is used to control the impact of gradient information. If this parameter is set to zero the proposed method reduces to traditional probability perturbation.

The use of gradual deformation to pick out good starting guesses for a deterministic history matching approach is also discussed. A sequential methodology

where gradual deformation is followed by a deterministic adjustment of the reservoir model is suggested. This method may be a good choice if the field exhibits smooth variations of permeability and if the distribution of permeability can be represented by a Gaussian distribution. However, the use of a deterministic method for the final adjustment may introduce geological artifacts which may ruin the geological consistency of the reservoir model.

Resume på dansk

Denne afhandling sammenfatter tre års arbejde med emnet statistiske metoder til history matching af olieproduktion. Arbejdet har resulteret i en omfattende beskrivelse af eksisterende teknikker og har desuden resulteret i udviklingen af en nyskabende metode, som kombinerer gradient information med en stokastisk history matching metode.

History matching er en vigtig del af driften af et oliefelt og er ofte forbundet med problemer relateret til kompleksiteten af reservoiret og selve størrelsen af reservoirsimuleringsmodellen. Begrebet *history matching* dækker over arbejdsprocessen, hvor de fysiske parametre i en reservoirsimuleringsmodel bliver justeret således, at en simulering af olieproduktionen stemmer overens med egentlige målte produktionsdata. For et producerende felt vil der generelt være målte productionsdata fra starten af produktionen indtil nutid. Disse målte data over tid kaldes historiske data. Modellen justeres, indtil reservoirsimuleringen giver produktionsdata, som er identiske med (eller tæt på) de historiske data. Reservoirmodellen skal først og fremmest bruges til at forudsige fremtidig produktion, hvilket er essentielt for mange dele af reservoirets drift og udvikling. Formålet med at history matche modellen er, at dens prædiktive egenskaber antages at være bedre, hvis modellen i det mindste beskriver historisk produktion. Med andre ord: Hvorfor skulle vi stole på prædiktioner fra en model, som ikke engang stemmer overens med historiske, målte data?

Denne afhandling er delt ind i tre dele, som er delt ind i et antal underkapitler:

- Baggrund og introduktion
  - Introduktion til history matching og litteraturstudie
  - Introduktion til geostatistik
  - Introduktion til reservoir simulering og adjoint sensitivitetsberegning

- History matching metoder
  - Gradual deformation
  - Probability perturbation
  - Deterministisk history matching
  - Hybrid metode
- Appendiks med uddybende detaljer vedrørende diverse emner

I første del introduceres grundbegreberne inden for geostatistik. Mange history matching metoder er baseret på et geostatistisk fundament, hvilket også gør sig gældende for de to stokastiske metoder omhandlet i denne afhandling. Derfor er en solid viden om geostatistik en forudsætning for at forstå disse metoder. En anden vigtig del af history matching er reservoir simulering. Den matematiske baggrund for strømning af fluider i porøse medier præsenteres sammen med en mere detaljeret gennemgang af den numeriske løsning af de diskretiserede ligninger. Både deterministiske metoder og den foreslåede hybrid metode fordrer en effektiv beregning af afledte af produktionsdata med hensyn til parametre i det diskretiserede grid, såsom permeaabilitet. Derfor behandles også en adjoint baseret tilgang til beregning af sådanne afledede. Detaljer relateret til implementeringen af adjoint metoden er angivet i appendiks.

Den del af afhandlingen, som omhandler metoder til history matching, er delt ind i fire kapitler, som hver især omhandler en specifik metode. I det sidste kapitel introduceres og diskuteres en nyskabende metode som kombinerer en eksisterende geostatistisk metode med gradient information.

Metoderne, som diskuteres, bliver i videst mulig udtrækning eksemplificeret ved anvendelse på history matching problemer. Et specielt fokus lægges på at eksemplificere brugen af den foreslåede hybridmetode, da denne er ny og derfor ikke er beskrevet i tidligere arbejder. Den hybride metode har vist sig at forbedre traditionel probability perturbation, idet hurtigere konvergens observeres. Arbejdet er begrænset til at omhandle binære reservoirmodeller, men den foreslåede metode burde kunne udvides til mere komplekse tilfælde med flere en to kategoriske variable. Den foreslåede metode forsøger at forbedre konvergensen af traditionel probability perturbation ved at inkludere kvalitativ gradient information. Denne information bliver ekstraheret fra gradienten af objektfunktionen ved anvendelse af et simpelt filter, som frafiltrerer numerisk små elementer i gradienten. En ny parameter, kaldet *degree of trust* introduceres. Denne parameter kan bruges til a styre indvirkningen af gradientinformationen. Sættes denne parameter til nul, reduceres den foreslåede metode til traditionel probability perturbation.

Brugen af gradual deformation metoden til at udvælge gode startgæt til en deterministisk history matching metode bliver også diskuteret. En sekventiel metode, hvor gradual deformation efterfølges af en deterministisk justering af reservoirmodellen foreslås. Denne metode kan være et godt valg, hvis reservoiret

udviser en jævn og glat variation af permeabiliteten, og hvis permeabiliteten kan repræsenteres af en normalfordelt distribution. Brugen af den deterministiske metode kan imidlertid introducere artifakter i den geologiske model, og den geologiske konsistens kan gå delvist tabt.

Contents

Part I

Background

Chapter 1

Introduction

A good description of key physical properties of an oil reservoir is essential for the reservoir engineer when the operation of the reservoir is planned. A long-standing problem in reservoir engineering is the inference of reservoir properties such as permeability or porosity from measured production data at the wells. The term history matching covers the process of adjusting a reservoir model such that a numerical simulation of the production agrees with actual measured production data, referred to as the *historical data*. Apart from honoring the historical data it is crucial that the reservoir model agrees with geological information. Such information may originate from core samples, seismic surveys, or experience from previous fields or outcrop data. The term *reservoir model* covers the geological description of a reservoir and includes many properties such as permeability, porosity, and many more. The best reservoir model will in general be the one that integrates most of the available knowledge. This knowledge may be very diverse ranging from historical production data to qualitative knowledge about the reservoir geology. The diversity of information sources makes it a difficult task to history match a reservoir model. The purpose of the present work is to investigate and device efficient history matching techniques which honor such diverse information sources.

The main reason for history matching of an oil or gas field is to use the reservoir model for prediction of future production. The main motive for history matching observed production data is that the predictive power of the resulting reservoir model is assumed to be best if at least the historic data is reproduced. A good prediction of the production is valuable when the operation of the field is planed and a properly history matched model may even be used to point out

areas in the reservoir with left-behind oil.

History matching is essentially a parameter estimation problem or an inverse problem where reservoir parameters are inferred from the measured production data. History matching goes under the category of *ill-posed* inverse problems because the problem is usually strongly underdetermined as a result of the large dimensionality of the unknown parameters compared to the number and the quality of the measurements. This also means that an infinite number of solutions to the problem exist, which all honor the measured data equally well. To deal with the ill-posedness of the problem, two approaches have gained a wide interest within the reservoir engineering community:

- Geostatistically based parameterizations
- Gradient-based optimization with regularization

Both approaches are discussed in this dissertation. Chapters 5 and 6 deal with two geostatistically based methods and Chapter 7 deals with gradient-based history matching and regularization. In Chapter 8 a hybridized method is presented. The hybrid method combines a geostatistical method with gradient information in order to yield a more efficient method.

1.1 Statement of the problem

The purpose of history matching is to formulate a reservoir model for which a numerical flow simulation will result in dynamic data which are consistent with actual observed data at the wells. Such dynamic data may be production data such as fluid flow rates, fraction of oil or gas in the production, and pressures in the wells. A common way of quantifying the mismatch between the simulated data and the observed data is to apply a sum-of-squares measure of the misfit:

$$E = \frac{1}{2} \sum_{i=1}^{N_{obs}} w_i (d_i^{obs} - d_i^{sim})^2, \tag{1.1}$$

where N_{obs} is the number of measurements of the production data. d_i is the ith measurement of production data. w_i are weighting factors which can be used to put special emphasis on particular data points.

With permeability as the unknown parameter the history matching problem can be stated as:

$$\mathbf{k}^* = \underset{\mathbf{k}}{\text{Argmin}} \left[E(\mathbf{k}) \right], \tag{1.2}$$

where $\mathbf{k}$ is the permeability tensor.

In realistic history matching problems, the unknown parameter is not restricted to grid block permeabilities but involves numerous other parameters

such as grid block porosities, relative permeabilities, initial saturations, fault locations, and more unknowns. Throughout this dissertation history matching will be restricted to inference of permeability. However, the adjoint code discussed in Chapter 4 has been extended such that derivatives with respect to porosity are available. This means that the deterministic method in Chapter 7 can be used for inference of porosity as well.

In field applications the size of the simulation grid may be of the order of $10^5 - 10^6$ grid blocks. Consequently, reservoir simulations take up a large amount of time and history matching can be a tedious task. Because of the computational load of reservoir simulation it is desirable to devise an efficient history matching technique which can integrate various geological information into the history match without running excessively many fluid flow simulations.

Traditionally, history matching has been carried out manually by experienced reservoir engineers who, from experience, know where and what to change in the reservoir model. The workflow of manual history matching is sketched in Figure 1.1. It may be difficult for the reservoir engineer to obtain a match of the production data and at the same time ensuring that the reservoir model is consistent with information regarding the geology of the reservoir. The work made in this project has focussed on the development and application of geostatistically based methods for history matching. The main strength of such methods is that the resulting reservoir model is easily constrained to statistical properties and other types of information regarding the geology.

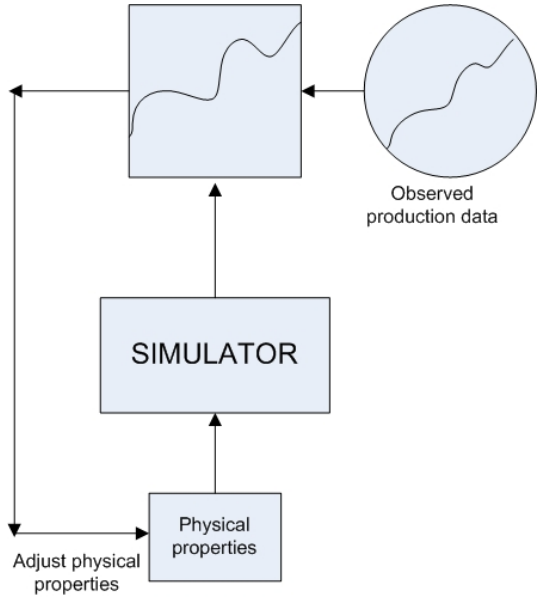


Figure 1.1: Principal workflow of manual history matching. Depending on the mismatch between simulated production data and observed data the reservoir engineer adjusts the reservoir model.

Chapter 2

Literature review

The present chapter presents a literature review of previous works related to history matching. Because of the importance of history matching the field has been subject to extensive research over the years. Consequently, this review will by no means be exhaustive but will focus on recent achievements and emphasis will be put on methods with a geostatistical framework.

Since history matching of production data is used to model the physical properties of the reservoir and to forecast future production there is a considerable economical motivation to improve the history match. Traditionally, history matching has been done manually by experienced reservoir engineers. Manual history matching is a tedious process which involves numerous reservoir simulations. Additionally, it may be difficult to integrate new information into the reservoir model. Consequently, a number of history matching techniques have been developed to facilitate the process of history matching. In this dissertation history matching techniques are grouped into two categories depending on the degree of human intervention needed to apply them:

Assisted history matching: Used to group techniques which require some expert intervention. This may be choices regarding the type of facies in certain areas or other decisions on the geology which have to be made during the course of the history matching.

Automated history matching: Denotes methods which can provide a history match automatically without any intervention from engineers or geologist during history matching. Decisions regarding the geology are specified prior to the history matching.

A method known as the pilot point method was introduced in de Marsily, Lavedan, Boucher & Fasanino (1984) who history matched pressure data from an interference test in a one-phase two-dimensional reservoir by adjusting the transmissibility field. An initial estimate of the permeability field (or transmissibility field) was found by kriging conditioned to hard data (i.e. samples). A number of grid points were then chosen as *pilot points* and used together with the sampled permeabilities to condition a kriged estimate. An objective function was optimized by adjusting the "measured" value at one or more pilot points. Thus, the pilot point method uses fictive measurements at pilot points as a parameterization of the history matching problem. The method is further developed in RamaRao, LaVenue, de Marsily & Marietta (1995) where gradient information is used to pick out the most influential pilot points. The pilot point method is in its original formulation an assisted history matching technique since the choice of pilot points is done manually. By inclusion of the gradient information the method can be automated because the gradient can be used to pick out suitable pilot points.

Another stochastic method was introduced in Hu (2000) and is denoted *gradual deformation*. The gradual deformation method exploits the fact that certain linear combinations of independent Gaussian fields conserve second order statistics. A number of parameters, known as *deformation parameters* are used to weigh each independent realization when the linear combination is formed. By adjusting these weights the resulting realization can be perturbed in order to achieve a better history match. In Hu (2002) a methodology is extended to handle dependent realizations. Gradual deformation is essentially a geostatistically based parameterization of the underlying optimization problem in history matching. The parameterization can reduce the mathematical complexity of history matching significantly. With gradual deformation the number of independent variables in the optimization problem can be reduced from possibly hundreds of thousands to only one. The gradual deformation method is a relatively simple methodology which makes it easy to implement in an existing history matching framework. However, the method is limited to deal with geologies which can be represented by a Gaussian distribution. Thus, some form of transformation of the adjusted parameter needs to be performed if it is not Gaussian. In practice, this means that gradual deformation is not suited for more complex geologies which cannot be derived from a Gaussian distribution (Caers 2007). The nature of the algorithm behind gradual deformation makes it possible to implement the gradual deformation method as a fully automated technique. In the literature as well as in this dissertation problems with convergence are reported. The gradual deformation method makes use of an extensive parameterization of the history matching problem which narrows down the search space used to minimize the production data mismatch. In order to improve convergence Hu & le Ravalec-Dupin (2004) incorporated gradient information into the gradual deformation method and improvements of the performance were reported. More information on gradual deformation is found in chapter 5 where the method is

applied to simple history matching problems.

A technique sharing some of the ideas behind gradual deformation was introduced by Caers (2003) and is referred to as the *probability perturbation method*. A more detailed description of the method is provided in Chapter 6. History matching of reservoirs where the geology is complex may be difficult to do manually. Also, application of gradient-based methods to categorical facies models is likely to challenge the effectiveness or even the applicability of a gradient-based method due to the discrete nature of the problem. The probability perturbation method (PPM) is highly flexible and works with Gaussian distributions and more complicated geological scenarios where multiple-point statistics (see section 3.5) is required to describe the geology sufficiently well. PPM exploits the algorithmic structure of sequential simulation algorithms which make use of local conditional probability functions. Through a conditional probability function PPM can perturb the probability of a given event on the grid block scale. Where gradual deformation perturbs physical properties at the grid block scale directly, PPM perturbs probabilities at the grid block scale which after a sequential simulation results in a perturbation of the desired physical properties. PPM does not rely on any assumptions regarding the statistics of the adjusted property, e.g. permeability. Therefore, the method is suited for any kind of geology which can be realized by sequential simulation.

Another technique which has been subject to extensive research in the recent 5 years is the so-called *ensemble Kalman filter*. The use of Kalman filters has a long tradition in the field of process control where the filter gives the optimal reconstruction of the states in a linear state space model (Madsen 1998). The governing flow equations for fluid flow in porous media do not result in a linear relationship between model parameters and production data (history matching is a non-linear problem). In order to account for the non-linearity an ensemble of Kalman filters are applied simultaneously to the history matching problem (Nævdal, Mannseth & Vefring 2002, Nævdal, Johnsen, Aanonsen & Vefring 2003). In Nævdal et al. (2002) permeability in a two phase, two-dimensional reservoir is history matched. The reservoir is operated by a smart well consisting of three compartments. The matched production data is well bottomhole pressure, pressure in the compartments, total fluid production, and the inflow of the two phases into each compartment. In Nævdal et al. (2003) the method is applied to a three phase problem. The Kalman filter updates dynamic properties like grid block pressures and saturations as well as grid block permeability (and other adjusted parameters) in each time step where a measurement is available. Thus, the reservoir model is not static but changes as more and more production data is integrated into the model. This may seem as an undesirable feature, since geological models of the reservoir traditionally are treated as static. A feature of Kalman filters is that integration of new production data does not require that the history matching is restarted from scratch. Ensemble Kalman filters represents a large area of research and it is out of the scope of this dissertation to describe the method in more detail.

Cheng, Kharghoria, He & Datta-Gupta (2004) use a streamline based method to compute sensitivities of watercut with respect to the permeability field of a real reservoir. Most optimization algorithms make use of the gradient of the objective function with respect to the adjustable variables. A major problem when standard finite difference simulators are used to simulate the production and calculate the respective sensitivities is the computational load associated with such operations. The faster streamline approach comes in very handy in this problem since this method is considerably faster. However, the streamlines are not able to account for what is referred to as *cross-streamline effects*. Such effects might be

- Mobility effects
- Rate changes
- Infill drillings

By the use of a hybrid method combining the streamline model with a commercial finite difference simulator (Eclipse) Cheng et al. (2004) have developed a fast gradient based history matching method which does not suffer from the inability to catch cross-streamline effects. The use of streamline derived sensitivities poses an unfortunate problem related to the principle of streamline simulators. A streamline derived sensitivity with respect to e.g. permeability only contains information averaged over the entire streamtube[1]. Therefore, a permeability change will be realized by multiplying the grid notes covered by the particular streamtube by a constant value. Such an update of the permeability field may easily be influenced by the distribution and shape of the streamlines. Consequently, geometric artifacts may arise during the history matching (Wang & Kovcek 2000, Agarwal & Blunt 2003).

More traditional optimization methods have also been applied to history matching. Wu, Reynolds & Oliver (1998) Li, Reynolds & Oliver (2003) use the Levenberg-Marquardt method to optimize the history match in a two-phase and a three-phase system, respectively. In these works an objective function of the following form is minimized:

$$E(\boldsymbol{m}) = \frac{1}{2}((\boldsymbol{m}-\boldsymbol{m}_{prior})^T \boldsymbol{C}_M^{-1}(\boldsymbol{m}-\boldsymbol{m}_{prior})+(\boldsymbol{g}(\boldsymbol{m})-\boldsymbol{d}_{obs})^T \boldsymbol{C}_D^{-1}(\boldsymbol{g}(\boldsymbol{m}))-\boldsymbol{d}_{obs}), \tag{2.1}$$

where $\boldsymbol{m}$ denotes the set of unknown reservoir parameters (vertical and horizontal permeability and well skin factors). Matrices $\boldsymbol{C}_M$ and $\boldsymbol{C}_D$ denote covariance matrices of the model parameters and observed data, respectively. The operator $\boldsymbol{g}$ denotes the process of running a reservoir simulation.

[1]The space between streamlines is referred to as a streamtube as it can be regarded as a separate tube decoupled from the rest of the reserovir.

Minimization of the objective function (2.1) is done by application of the Levenberg-Marquardt algorithm. An adjoint methodology is used to calculate the sensitivities needed for the optimization. The described methodology is not based on a geostatistical framework and therefore the result is not guaranteed to be consistent with geological information such as semivariograms. However, the objective function as given in equation (2.1) can be considered as a regularized form of the general measure of production data misfit given in equation (1.1). The objective function penalizes deviations from a prior model which can be seen as a regularization of the problem. The regularizing term ensures some consistency of the result and is a convenient means of keeping a handle on the optimizer to avoid unphysical or highly implausible results. The use of regularization techniques is discussed in Chapter 7.

2.1 Summary

The existing literature contains many works related to the subject *history matching*. Many proposed methods for automated or assisted history matching are based on a geostatistical framework. This includes the gradual deformation method which reduces the problem of history matching to a relatively simple optimization problem. The use of the gradual deformation method requires that the reservoir geology can be represented by a Gaussian distribution. This means that complex reservoir geologies with curvilinear features are not suited for history matching with the gradual deformation method. Another method, denoted *probability perturbation*, can be applied to history matching of fields with such complex geological features. Common to gradual deformation and probability perturbation is that the convergence of the methods is slow because of their stochastic nature. The parameterizations used in the methods restricts the search space available for optimization resulting in poor convergence. The gradual deformation method has been extended to use gradient information in order to improve the convergence. Such an extension has until now not been presented for the probability perturbation method.

Chapter 3

Geostatistics

This chapter deals with the mathematical, statistical, and physical background of geostatistics. The motivation for this chapter is to provide a general knowledge on some of the geostatistical terms used in the dissertation. The part related to geostatistics relies on the excellent discussion on geostatistics provided by Goovaerts (1997).

First, the basics of classical geostatistics are discussed followed by a short introduction to multiple-point statistics. The discussion is supplemented with examples and figures where it is possible.

3.1 Basic geostatistics

Geostatistics deals with spatially varying properties and has found a wide range of applications, from mining and oil exploration to agricultural purposes as well as image analysis. The concepts in geostatistics can be applied to many problems which exhibit a spatial dependence. Originally, geostatistics was developed for mining and mineral exploration in the 1950's and was further developed by the French engineer Georges Matheron during the 1960's. Traditionally, geostatistics has been limited to second-order statistics such as covariance or variograms. However, during the last ten years the more advanced multiple-point statistics has gained attention. Multiple-point statistics has enabled geostatisticians to deal with complex geologies which are difficult or impossible to describe with the traditional 2-point statistics used in the "classical" geostatistics.

First, the basic statistical properties are introduced. Consider some property

z which is assumed to be a spatially distributed variable (for instance porosity or permeability). $z(r)$ is now treated a realization of the stochastic variable $Z(r)$. r is indicating that z is a regionalized variable. The expectational value of $Z(r)$ is given as

$$\mathrm{E}\{Z(r)\} = \mu(r), \tag{3.1}$$

and the covariance is given by

$$C(r,h) = \mathrm{E}\{[Z(r) - \mu(r)][Z(r+h) - \mu(r)]\}. \tag{3.2}$$

If the mean of the property in question is independent of position, that is $\mu(r) = \mu$, Z is denoted as 1th order stationary. If also the covariance is independent of the position, Z is denoted as 2nd order stationary.

The variogram is defined as

$$2\gamma(h) = \mathrm{E}\{[Z(r) - Z(r+h)]^2\}. \tag{3.3}$$

The intrinsic hypothesis of geostatistics states:

Hypothesis The variogram is independent of the position and only depends on the lag, i.e. the distance between the two points (Goovaerts 1997).

Mathematically stated the intrinsic hypothesis says

$$\gamma(r,h) = \gamma(h). \tag{3.4}$$

If the stochastic variable Z is second order stationary the intrinsic hypothesis is valid.

Under 2nd order stationarity the semivariogram is given as

$$\gamma(h) = C(0) - C(h). \tag{3.5}$$

The validity of equation (3.5) can be shown by the following proof:

The variance is defined as

$$\sigma^2 = \mathrm{E}\{[Z(h) - \mu]^2\} = \mathrm{E}\{Z^2(h)\} - \mu^2, \tag{3.6}$$

and the covariance is defined as

$$\begin{aligned} \mathrm{C}(h) &= \mathrm{E}\{[Z(r) - \mu][Z(r+h) - \mu]\} \\ &= \mathrm{E}\{Z(r) \cdot Z(r+h)\} - \mu\mathrm{E}\{Z(r)\} - \mu\mathrm{E}\{Z(r+h)\} + \mu^2 \\ &= \mathrm{E}\{Z(r) \cdot Z(r+h)\} - \mu^2. \end{aligned} \tag{3.7}$$

The variogram is defined as

$$2\gamma(h) = \mathrm{E}\{[Z(r) - Z(r+h)]^2\} = \mathrm{E}\{Z^2(r)\} + \mathrm{E}\{Z^2(r+h)\} - 2\mathrm{E}\{Z(r) \cdot Z(r+h)\}$$

From equations (3.6) and (3.7) the expression becomes

$$2\gamma(h) = 2(\sigma^2 + \mu^2) - 2(\mathrm{C}(h) + \mu^2) \Leftrightarrow \gamma(h) = \sigma^2 - \mathrm{C}(h) = \mathrm{C}(0) - \mathrm{C}(h).$$

From a number of observations the experimental semivariogram is easily calculated by the formula

$$\hat{\gamma}(h) = \frac{1}{2N(h)} \sum_{k=1}^{N(h)} \Big(z(r_k) - z(r_k + h) \Big)^2, \tag{3.8}$$

where $N(h)$ is the number of pairs of observations with a distance corresponding to the lag h.

A number of empirical semivariogram models can be used to describe spatial variation of z in the particular space, e.g. the oil reservoir. In equations (3.9a - 3.9c) the three most common models are given.

Spherical model

$$\gamma^*(h) = \begin{cases} 0 & h = 0 \\ C_0 + C_1 \left(\frac{2}{3} \frac{h}{R} - \frac{1}{2} \frac{h^3}{R^3} \right) & 0 < h < R \\ C_0 + C_1 & h \geq R \end{cases} \tag{3.9a}$$

Exponential model

$$\gamma^*(h) = \begin{cases} 0 & h = 0 \\ C_0 + C_1 \left(1 - \exp\left(-\frac{3h}{R} \right) \right) & h > 0 \end{cases} \tag{3.9b}$$

Gaussian model

$$\gamma^*(h) = \begin{cases} 0 & h = 0 \\ C_0 + C_1 \left(1 - \exp\left(-\frac{3h^2}{R^2} \right) \right) & h > 0 \end{cases} \tag{3.9c}$$

The parameter C_0 is denoted the *nugget effect* and acts as a measure of measurement error and discontinuity in the vicinity of $h = 0$. The quantity $C_0 + C_1$ is denoted *the sill* which is the value of the semivariogram as the lag tends to infinity. R is the *range of influence* and is a measure of the extent to which a measurement has an impact on a new estimation.

3.2 Spatial interpolation

The most simple way of approximating a property field from a set of spatially distributed measurements of the property in question is to approximate the field by the mean value at all positions. This, however, does not honor any variability detected through the measurements, nor does it agree with the experiences made in practice in geostatistics or reservoir engineering. As a consequence, one might carry out a spatial interpolation where the spatial distance from a measurement to the point of estimation acts as a weighting factor. This can be achieved by weighting each measurement with its inverse distance to the estimated point. Equations 3.10 and 3.11 shows an estimation scheme with distance weighting. d_i is the distance from the estimation point to the ith sample and N is the number of samples (measurements).

$$w_i = \frac{1/d_i}{\sum_{j=1}^{N} 1/d_j} \tag{3.10}$$

$$\hat{z} = \boldsymbol{w}\boldsymbol{z}, \quad \text{where } \boldsymbol{z} = [z(r_1), z(r_2)...z(r_N)] \tag{3.11}$$

A distance weighting as described above is a huge improvement with respect to its applicability compared to just using the mean of the samples as an estimation. However, the approach does not honor any knowledge about the statistics of the field. This includes knowledge about the variance in the form of a semivariogram or semivariogram model. A technique which can incorporate such knowledge into the estimation procedure is *kriging*.

3.2.1 Kriging

The background of kriging methods is presented in the following section. Again, a linear estimator for the property z is the basis for the method. In its general form the linear estimator takes the form

$$\hat{z}_0 = w_0 + \sum_{i=1}^{N} w_i z_i = w_0 + \boldsymbol{w}^T \boldsymbol{z}, \tag{3.12}$$

where w_i is the weight of the ith sample, z_i. The z_i's are assumed to be realizations of the stochastic variable Z_i consisting of two parts, a mean value and a residual - i.e. $Z_i = \mu_i + \epsilon_i$. The variance of the residuals is assumed to have a constant variance of σ^2 and a mean of zero. The expected value of the estimation error is

$$\mathrm{E}\{Z_0 - \hat{Z}_0\} = \mathrm{E}\{Z_0 - w_0 - \boldsymbol{w}^T \boldsymbol{Z}\} = \mu_0 - w_0 - \boldsymbol{w}^T \boldsymbol{\mu}, \tag{3.13}$$

where $\boldsymbol{\mu} = [\mu_1\ \mu_2...\mu_N]^T$.

A central estimate implies that the expected estimation error is zero which implies that equation (3.13) should be equated to 0:

$$\mu_0 - w_0 - \boldsymbol{w}^T \boldsymbol{\mu} = 0. \tag{3.14}$$

The estimation variance is given by (for explanation on the algebraic operations refer to any book on multivariate statistics, e.g. (Conradsen 2003)).

$$\begin{aligned}\sigma_E^2 = \mathrm{Var}\{Z_0 - \hat{Z}_0\} = \mathrm{Var}\{Z_0\} + \mathrm{Var}\{w_0 + \boldsymbol{w}^T \boldsymbol{Z}\} - 2\mathrm{Cov}\{Z_0, w_0 + \boldsymbol{w}^T \boldsymbol{Z}\} \\ = \sigma^2 + \boldsymbol{w}^T(\boldsymbol{C}\boldsymbol{w} - 2\mathrm{Cov}\{Z_0, \boldsymbol{Z}\}).\end{aligned} \tag{3.15}$$

$\boldsymbol{C}$ is the covariance matrix of $\boldsymbol{Z}$.

Simple Kriging

In simple kriging the estimation variance is minimized. From equations (3.12) and (3.14) the following expression is obtained

$$\hat{Z}_0 - \mu_0 = \boldsymbol{w}^T(\boldsymbol{Z} - \boldsymbol{\mu}). \tag{3.16}$$

The partial derivative of the estimation variance with respect to the weight vector is equated to zero:

$$\frac{\partial \sigma_E^2}{\partial \boldsymbol{w}} = 2\boldsymbol{C}\boldsymbol{w} - 2\mathrm{Cov}\{Z_0, \boldsymbol{Z}\} = \boldsymbol{0}. \tag{3.17}$$

This gives a system of equations for simple kriging:

$$\begin{aligned}\boldsymbol{C}\boldsymbol{w} = \mathrm{Cov}\{Z_0, \boldsymbol{Z}\} \Leftrightarrow \\ \begin{bmatrix} C_{11} & \dots & C_{1N} \\ \vdots & \ddots & \vdots \\ C_{N1} & \dots & C_{NN} \end{bmatrix} \begin{bmatrix} w_1 \\ \vdots \\ w_N \end{bmatrix} = \begin{bmatrix} C_{01} \\ \vdots \\ C_{0N} \end{bmatrix},\end{aligned} \tag{3.18}$$

where C_{ij} is the covariance between the sampled points and where C_{0i} is the covariance between the sampled data and the estimated point. The covariances can be found from the semivariogram. The variance of the estimation is

$$\sigma_{SK}^2 = \sigma + \boldsymbol{w}^T(\boldsymbol{C}\boldsymbol{w} - 2\mathrm{Cov}\{Z_0, \boldsymbol{Z}\}) = \sigma - \boldsymbol{w}^T \mathrm{Cov}\{Z_0, \boldsymbol{Z}\}. \tag{3.19}$$

Ordinary Kriging

In ordinary kriging the mean value of Z_i is assumed to be constant for all measurements. This results in the expression

$$\mathrm{E}\{Z_0 - \hat{Z}_0\} = \mu_0(1 - \boldsymbol{w}^T\mathbf{1}) - w_0, \quad \text{where } \mathbf{1} = [1\ 1...1]. \tag{3.20}$$

Equation (3.20) only holds (for all μ_0) when $w_0 = 0$ and $1 - \boldsymbol{w}^T\mathbf{1} = 0$. The minimization of the estimation variance is now carried out under the constraint $1 - \boldsymbol{w}^T\mathbf{1} = 0$. In practice, such a constraint is included through a Lagrange multiplier (2λ) . The constraint is added to the objective function (3.15) giving the Lagrange function

$$\mathcal{L} = \sigma_E^2 + 2\lambda(\boldsymbol{w}^T\mathbf{1} - 1). \tag{3.21}$$

The partial derivatives with respect to $\boldsymbol{w}$ and λ are equated to 0

$$\frac{\partial \mathcal{L}}{\partial w} = 2\boldsymbol{C}\boldsymbol{w} - 2\mathrm{Cov}\{Z_0, \boldsymbol{Z}\} + 2\lambda\mathbf{1} = 0 \tag{3.22a}$$

$$\frac{\partial \mathcal{L}}{\partial \lambda} = 2(\boldsymbol{w}^T\mathbf{1} - 1) = 0. \tag{3.22b}$$

The system of equations for ordinary kriging is thus given as

$$\begin{aligned} \boldsymbol{C}\boldsymbol{w} + \lambda\mathbf{1} &= \mathrm{Cov}\{Z_0, \boldsymbol{Z}\} \\ \mathbf{1}^T\boldsymbol{w} &= 1. \end{aligned} \tag{3.23}$$

More precisely, the equations take the form

$$\begin{bmatrix} C_{11} & \dots & C_{1N} & 1 \\ \vdots & \ddots & \vdots & \vdots \\ C_{N1} & \dots & C_{NN} & 1 \\ 1 & \dots & 1 & 0 \end{bmatrix} \begin{bmatrix} w_1 \\ \vdots \\ w_N \\ \lambda \end{bmatrix} = \begin{bmatrix} C_{01} \\ \vdots \\ C_{0N} \\ 1 \end{bmatrix}. \tag{3.24}$$

Again, the covariances are obtained from the semivariogram, which may be approximated by experimental data or by known or expected values of the sill, range of influence, etc..

The variance associated with ordinary kriging is given as

$$\sigma_{OK}^2 = \sigma + \boldsymbol{w}^T(\boldsymbol{C}\boldsymbol{w} - 2\mathrm{Cov}\{Z_0, \boldsymbol{Z}\}) = \sigma - \boldsymbol{w}^T\mathrm{Cov}\{Z_0, \boldsymbol{Z}\} - \lambda. \tag{3.25}$$

Application of kriging to clustered or non-equidistant samples may lead to negative weights and maybe even negative estimates of permeability, porosity, etc.. A number of techniques can be implemented to circumvent this rather

disappointing feature of kriging (da Rocha & Yamamoto 2000). Se also section 3.3 which discusses declustering techniques. An important detail lies in the fact that the kriging variance does not depend on the values of the samples. Only, the distribution of the samples has an impact on the kriging variance.

3.2.2 Accounting for anisotropy

In the previous sections a statistically homogeneous two-dimensional field has been the basis for a linear estimation of permeability. Such a field is called *isotropic* indicating that the variability of the property in question (for instance permeability) does not depend on the direction of the separation vector h - only the length is responsible for any variability. In many applications such an assumption is invalid. An oil reservoir may be systematically build up by areas with unique properties according to some underlying geologic process.

Hohn (1999, Ch. 2) defines *geometric anisotropy* as the situation where the sill is independent of the direction of $\boldsymbol{h}$ and only the range of influence depends on the direction.

Consider the two spherical semivariograms which account for horizontal variability and vertical variability, respectively:

$$\gamma_1(h) = \frac{3h}{2R_1} - \frac{h^3}{2R_1^3} \tag{3.26}$$

$$\gamma_2(h) = \frac{3h}{2R_2} - \frac{h^3}{2R_2^3}. \tag{3.27}$$

The *anisotropy ratio* is defined as

$$\nu = \frac{R_1}{R_2}. \tag{3.28}$$

Inserting νh in the first semivariogram gives

$$\gamma_1(\nu h) = \frac{3R_1 h}{2R_2 R_1} - \frac{R_1^3 h^3}{2R_2^3 R_1^3} = \frac{3h}{2R_2} - \frac{h^3}{2R_2^3} = \gamma_2(h). \tag{3.29}$$

Thus, the semivariogram in the vertical direction can be calculated by the semivariogram for the horizontal direction if the separation distance is multiplied with the anisotropy ratio. Such an approach, however, only enables the user to determine the semivariogram in two directions - the horizontal and the vertical directions. If the direction lies between these a change of coordinate system must be applied. Let a linear coordinate transformation be given as

$$\begin{aligned} h'_h &= a_{11} h_h + a_{12} h_v \\ h'_v &= a_{21} h_h + a_{22} h_v, \end{aligned} \tag{3.30}$$

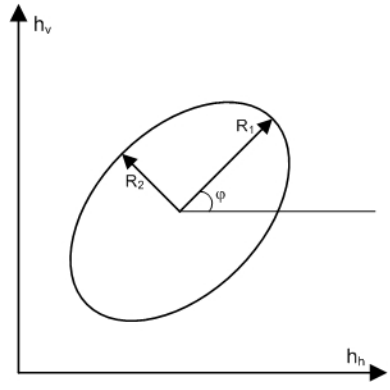


Figure 3.1: Illustration of the coordinate transformation.

where $\boldsymbol{h} = [h_h\ h_v]^T$ and $\boldsymbol{h}' = [h'_h\ h'_v]^T$.

Figure 3.1 shows an ellipsoid with axes corresponding to the directions of influences associated with the ranges R_1 and R_2. In the following discussion ϕ will denote the angle between the principal axis of the ellipsoid and the horizontal axis as depicted on Figure 3.1.

Hohn (1999, Ch. 2) gives the transformation

$$\boldsymbol{h}' = \begin{bmatrix} 1 & 0 \\ 0 & \nu \end{bmatrix} \begin{bmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{bmatrix} \boldsymbol{h} = \boldsymbol{SRh}. \tag{3.31}$$

The transformation (3.31) aligns the kriging coordinate system with the axes of influence.

The semivariogram value can now be determined by the use of the transformed coordinates:

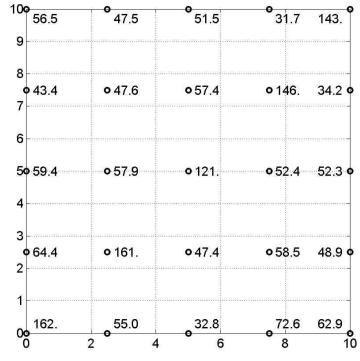
$$\gamma(h) = \frac{3\|h'\|}{2R_1} - \frac{\|h'\|^3}{2R_1^3}, \tag{3.32}$$

where

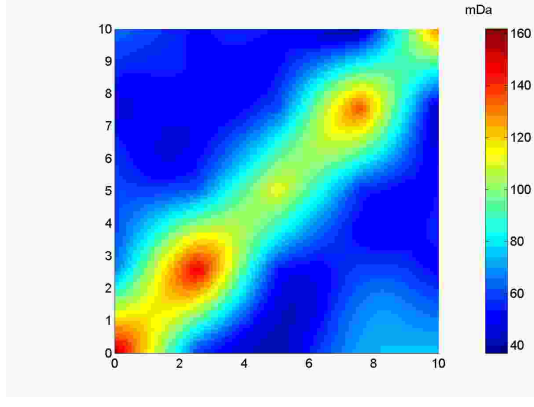
$$\|h'\| = \sqrt{h'^2_h + h'^2_v}. \tag{3.33}$$

Figure 3.2(a) shows a quadratic field where samples of permeability have been made in an ordered grid. The samples reveal that a high-permeable area is located on the diagonal going from the lower left corner to the upper right corner. Figure 3.2(b) shows a kriged permeability field based on the samples. The kriged field appears to be smooth and does not seem to represent a realistic geology. This smoothing effect is a general feature of kriging and results from the minimization of the error variance. This has lead to the development of other techniques to generate more realistic property fields, e.g. the sequential

methods described in section 3.4. The kriged result minimizes the estimation variance and is therefore unique, hence kriging does not allow generation of multiple realizations.



(a) Samples.



(b) Kriged field.

Figure 3.2: Illustration of a kriging result based on the samples shown to the left. The kriged permeability field appears smooth - a general feature of kriging.

3.3 Data declustering

If the measured data from the oil field is not evenly distributed over the entire reservoir area the geologist or the reservoir engineer is in danger of making erroneous conclusions about the statistics of the reservoir. Figure 3.3 shows two

quadratic areas in which samples have been made. In the left area 4 samples are clustered closely in the lower left corner. In the right area the samples are more evenly distributed. Note that both areas share two sample points. The estimated means based on the 5 samples in each area are 8.2 and 3.7, respectively. Intuitively, the lower mean seems most reasonable because the sampling in the right area appears to be more representative for the whole area. When an arithmetic mean is calculated from the samples in the left area the high-valued area in the left corner is overrepresented resulting in a high estimated mean. When the data is clustered it is desirable to be able to put more weight on measurements which are placed in the less densely sampled areas. This can be achieved by the use of certain *declustering* techniques.

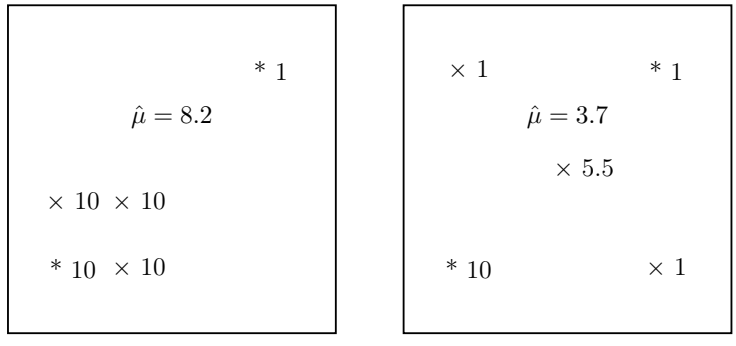


Figure 3.3: Example showing how a clustered data set may lead to erroneous conclusions about the statistics of a spatially varying variable. Common sample points are marked with *.

3.3.1 Common techniques for data declustering

Isaaks & Srivastava (1989), Goovaerts (1997), and Deutsch & Journel (1998) contain discussions on techniques used to decluster data sets. Especially the first reference gives a good introduction to the concept of declustering. The two main methods are *the polygonal method* and *the cell-declustering technique.*

In the polygonal method each sample point is assigned an area corresponding to a polygonal surrounding the sample point with edges placed halfway between the point in question and the other points - se Figure 3.4. This means that the area of the polygon corresponds to the set of points which is closest to the particular sample point. Weighting the individual points with weights proportional to the surface area of the polygon will ensure that the impact of samples from densely sampled areas is reduced. Thus, the mean and variance may be

expressed as:

$$\hat{\mu} = \frac{1}{\mathcal{A}} \sum_{\alpha=1}^{n} \omega_\alpha z(\boldsymbol{u}_\alpha) \tag{3.34a}$$

$$\hat{\sigma}^2 = \frac{1}{\mathcal{A}} \sum_{\alpha=1}^{n} \omega_\alpha [z(z(\boldsymbol{u}_\alpha) - \hat{\mu})]^2, \tag{3.34b}$$

where

$$\mathcal{A} = \sum_{\alpha=1}^{n} \omega_\alpha = \text{Total Area}. \tag{3.34c}$$

ω_α denotes the surface area of particular polygons.

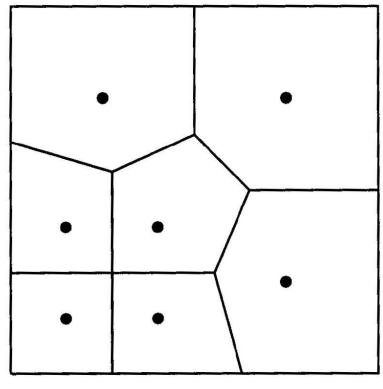


Figure 3.4: Sketch showing samples and their corresponding polygonals of influence. Figure taken from Goovaerts (1997).

The most outspoken difference between the cell-declustering technique and the polygonal method is that the division of the area into subspaces of polygonals is unique - the cell-declustering technique on the other hand can result in several outcomes. The idea behind this technique is to divide the area into a number of equally sized rectangular subspaces (cells) - se Figure 3.5. Then the number, B, of non-empty cells is counted. The number of samples within the B cells are counted - the quantity $n_b, b = 1, ..., B$ refers to these numbers.

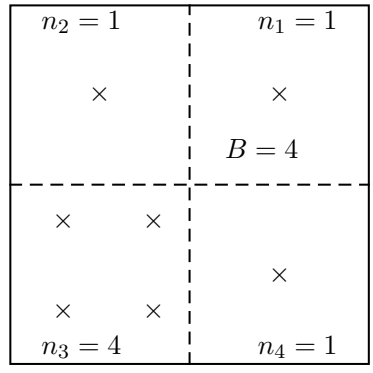


Figure 3.5: Illustration of the cell-declustering technique.

The mean and variance is then given by:

$$\hat{\mu} = \sum_{\alpha=1}^{n} \omega_\alpha z(\boldsymbol{u}_\alpha) \tag{3.35a}$$

$$\hat{\sigma}^2 = \sum_{\alpha=1}^{n} \omega_\alpha [z(z(\boldsymbol{u}_\alpha) - \hat{\mu})]^2 \tag{3.35b}$$

$$\omega_\alpha = \frac{1}{B \cdot n_b}. \tag{3.35c}$$

Thus, samples from cells with many samples are made less dominant than samples from cells with few samples.

3.4 Sequential Simulation

The previously discussed kriging methods can be generalized to be part of the so-called *estimation methods*. These methods are characteristic in the sense that the estimated field is unique because the estimation variance is minimized. This means that whenever a number of sample values are present and the variability (the semi-variogram) is determined, only *one* realization will fulfill the requirement of minimum estimation variance. The so-called *simulation techniques* are different with respect to this property. By relaxing the requirement of minimum estimation variance the simulation techniques are able to provide *multiple* and *equiprobable* realizations. One of the frequently used simulation methods is the *sequential Gaussian simulation* method. Sequential Gaussian simulation relies on the assumption that the distribution of the simulated property locally can be

described by a Gaussian distribution. During the past decade a number of new simulation techniques have appeared and made it possible to simulate complex geological scenarios (Strebelle 2000).

Whereas kriging can be done for all unknown points simultaneously, the simulation techniques visit each unknown point sequentially. Figure 3.6 illustrates the workflow in sequential simulation.

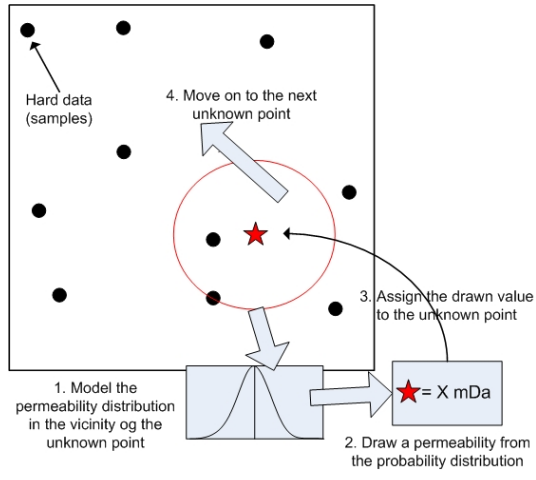


Figure 3.6: Illustration of the workflow in sequential simulation. The way the local probability function is modelled varies from method to method. In Gaussian simulation the pdf is modelled as Gaussian.

A sequential simulation can be summarized by the following algorithm:

1. Model the cumulative density function (ccdf) at the first location $\boldsymbol{u}_1$ conditional on the original n samples:

$$F(\boldsymbol{u}_1; z|(n)) = \text{Prob}\{Z(\boldsymbol{u}_1) \leq z|(n)\}$$

2. Draw a value from the ccdf

3. At position $\boldsymbol{u}_i$ model the ccdf conditional on the samples as well as previously simulated (drawn) values:

$$F(\boldsymbol{u}_i; z|(n+i-1)) = \text{Prob}\{Z(\boldsymbol{u}_i) \leq z|(n+i-1)\}$$

4. Draw the ith value, set $i = i + 1$

5. Repeat steps 3 and 4 until all grid notes have been visited

In sequential Gaussian simulation the simulation of a new value from the local conditional distribution amounts to solving a local kriging system. Thus, points 2 and 4 in the algorithm outline correspond to solving a local kriging problem.

If the unknown points are visited in a too systematic manner there will be a risk of creating geologic artifacts. This problem arises because previously simulated notes are used to model future pdfs in their neighborhood. Consequently, visiting is usually done in a random or pseudo-random order to avoid creation of geologic artifacts. The order is usually controlled by a seed number specified by the user. By changing the seed multiple equiprobable realizations can be obtained.

Figure 3.7 shows a sequential Gaussian simulation of a permeability field conditioned to the samples in Figure 3.2(a). Compared to the kriged result the simulated field is less smooth and appears more realistic.

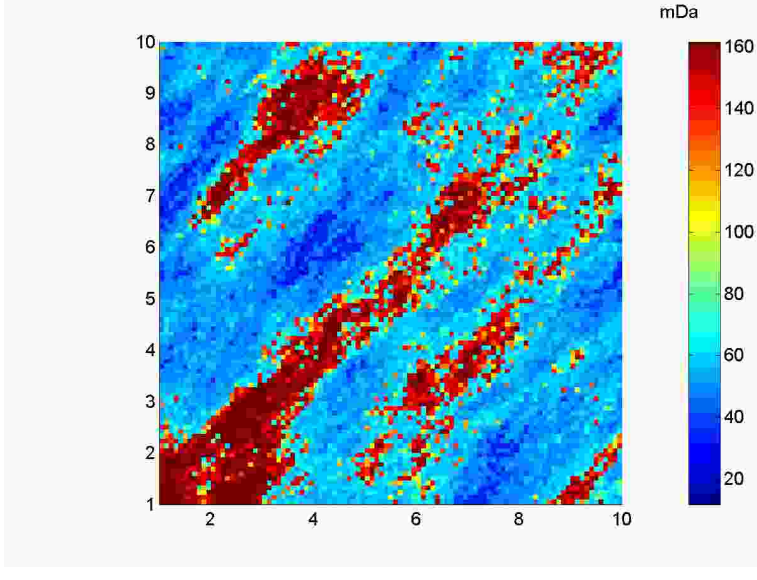


Figure 3.7: Sequential Gaussian simulation of a permeability field based on the samples shown in Figure 3.2(a). Compared to the kriged result the simulated field is less smooth and appears more realistic.

The term "sequential simulation" is used to describe methods which make use of the sequential algorithm described previously. The difference between various sequential simulation algorithms lies in how the local conditional distribution function is modelled.

3.5 Multiple-point statistics

Kriging and sequential Gaussian simulation are said to honor two-point statistics since both methods only involve semivariograms and no other statistical moments. However, 2-point statistics are not sufficient to reproduce curvi-linear features such as channels or other geological objects (Caers 2003, Strebelle 2000, Strebelle 2002). As previously shown, covariance is calculated from a 2-point templates:

$$C(\boldsymbol{h}) = \mathrm{E}\{Z(\boldsymbol{u})Z(\boldsymbol{u}+\boldsymbol{h})\}. \tag{3.36}$$

$(n+1)$-point statistics are given as ((Journel 2006)) :

$$C(\boldsymbol{h}_1, \boldsymbol{h}_2, .., \boldsymbol{h}_n) = \mathrm{E}\{Z(\boldsymbol{u}) \prod_{i=1}^{n} Z(\boldsymbol{u}+\boldsymbol{h}_i)\}. \tag{3.37}$$

Thus, an (n+1)-point template is needed for (n+1)-statistics. Representation of complex geological objects requires higher order statistics than semivariograms. However, reproduction of higher order statistics is a tedious task and many of the preliminary implementations of algorithms which took multiple-point statistics into account were CPU and RAM demanding (Strebelle 2002). One approach to simulation of complex geological geometries is object based simulation where a number of predefined objects are assigned to the grid blocks during simulation. A drawback of these methods is that local conditioning to data is difficult and it may be hard to represent complex geologies with a limited number of objects (Strebelle 2002).

The most effective and reliable approaches to simulation of complex geological geometries are the training image based methods. These methods require a so-called *training image* (TI) which represents the type of geology to be simulated. Multiple-point statistics are extracted from the TI in various ways depending on the individual methodology. Strebelle (2000) presented Snesim, a sequential simulation method based on the concept of training images. In a preprocessing step Snesim scans the TI and builds a search tree of events. During simulation a search through this tree is made when the local conditional distribution function is computed. The strength of this approach is that the method is highly flexible, i.e. simulation of a new type of geology only requires that a new TI is constructed in accordance with the desired geology. Thus, Snesim makes it possible to generate multiple realizations based on conceptual models of the geology. Such conceptual models may be supplied from geologists, seismic, outcrops, or other sources. Snesim is only applicable to categorical (indicator) variables. In the paper of Zhang, Bombarde, Strebelle & Oatney (2006) an alternative to Snesim is introduced. The described method is named Filtersim because it uses filters to classify different patterns from the training image. This method works with continuous variables and categorical variables. In Filtersim the data events found by scanning the training image are grouped under

a number of class prototypes. Via a distance measure the constellation of points in the vicinity of the simulated point is compared to the class prototypes. From the group of patterns belonging to the prototype with the largest similarity with the constellation a pattern is drawn and attached to the simulated point.

Figure 3.8 shows an example of a TI with a channel structure. The TI contains binary facies indicators for channel facies and non-channel facies. Figures 3.9(a and b) show two realizations of facies conditioned to the TI from Snesim. The Snesim results reproduce the channel structure in both realizations but show a large variation mutually.

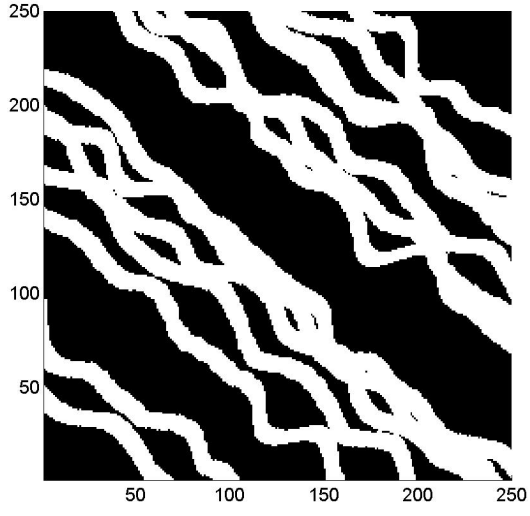
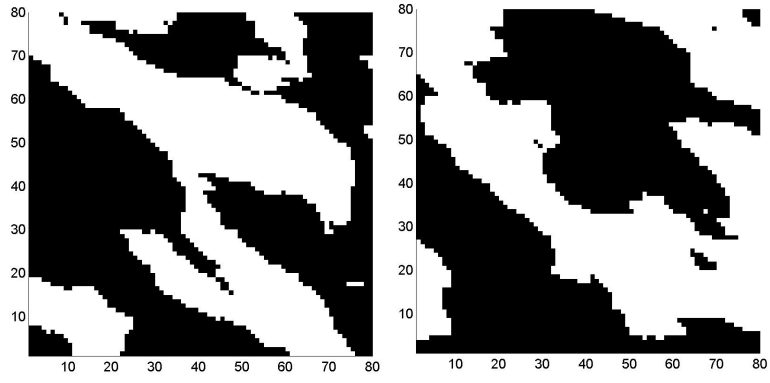


Figure 3.8: Training image with a channel structure.

3.6 Summary of geostatistics

The basics of geostatistical methods such as kriging, sequential simulation, and multiple point statistics have been presented. In later chapters these concepts will make up the framework for history matching techniques. The probability perturbation method which is discussed in Chapters 6 and 7 exploits the algorithmic structure of sequential simulation and the methodology of this method should be more clear after the introduction to sequential simulation provided in the present chapter. The Snesim algorithm has been chosen as the sequential simulator for generation of realizations conditioned to multiple-point statistics in this dissertation. The Snesim algorithm is easy to integrate into a history matching procedure and the algorithm is available through the SGeMS software



(a) First realization. (b) Second realization.

Figure 3.9: Two simulation results from Snesim conditioned to the geology depicted in the training image shown in Figure 3.8.

package from Stanford Center for Reservoir Forecasting (Remy 2004).

Chapter 4

Reservoir simulation

This section deals with reservoir simulation with the main purpose of introducing important concepts related to simulation of fluid flow in porous media which is an integral part of history matching. An in-house reservoir simulator, denoted BOSS, has been the basis of the present work. This simulator was developed at Center for Phase Equilibria and Separation Precesses (IVC-SEP) at the Technical University of Denmark by Morten Rode Kristensen and the author. This section introduces the basic concepts of fluid flow in porous media as well as the numerical methods used to solve the discretized model equations. For a comprehensive discussion of these topics and a detailed description of the simulator the reader may consult Johansen & Kristensen (2006).

History matching is a computationally expensive process mainly due to the computational load of reservoir simulation. In certain processes, e.g. gas injection, an accurate description of the thermodynamical behavior of the reservoir fluids is necessary to capture important physical mechanisms. For instance, in gas injection processes it is important that the solubility of the injected gas in the oil phase is properly modeled and accounted for. In such cases it is important that the oil and gas phases are treated as complex mixtures of given components or pseudo-components when the fluid flow is simulated. When the reservoir simulator treats the fluids as complex mixtures it is said to be *compositional*. Even with today's modern supercomputers compositional reservoir simulations are considered to be very heavy. Consequently, compositional simulation is often carried out on coarse simulation grids in order to limit the time of computation (Barker & Fayers 1994). Reservoir engineers have therefore resorted to simplifying the physics of the fluids by introducing the so-called *black-oil* description of

the fluid phases. Black-oil simulators treat the oil component as a single component fluid. Analogously, the gas component is treated as a single component fluid. Thus, a black-oil system contains three components:

- water
- oil
- gas

The characteristic assumptions in black-oil simulators are:

- Oil and water components are immiscible
- Gas is dissolvable in both oil and water phases
- Oil and water does not vaporize into the gas phase

In addition the reservoir is treated as isothermal.

Black-oil simulations are less demanding with respect to computation time compared to compositional simulation. Of course, the trade-off is that the physics are treated less realistically. Because of the reduced computational work related to the black-oil formulation this approach is relevant for history matching purposes. In the present work a simplified two-phase system makes out the basis for reservoir simulation. The phases will be treated as immiscible. Simplification is done in order to reduce computational work and it is deemed that the applicability of the history matching techniques treated later in the dissertation is demonstrated sufficiently well with the simplified description of the reservoir fluids.

4.1 Governing equations

The governing flow equations for flow of a two-phase system (water and oil) in heterogenous porous media are now derived. The porous media belongs to the domain Ω and has the boundary Γ. It is assumed that the temperature is constant over the domain Ω and that the fluids do not flow across the boundaries. As mentioned above miscibility between phases is not accounted for. The phases are assumed compressible and the reservoir rock is also treated as compressible in order to account for porosity changes with varying pressures. The void space in the reservoir rock is always filed with fluids which translates into the following basic relationship:

$$S_w + S_o = 1, \tag{4.1}$$

where S_w and S_o are water saturation and oil saturation, respectively.

The governing flow equations are derived from considerations of mass conservation. In the following treatment capillary effects are neglected. If dispersion and diffusion effects are neglected the governing flow equations for water and oil are (Aziz & Settari 1979):

$$\frac{\partial}{\partial t}\left[\varphi\rho_w S_w\right] - \nabla\cdot\left[\rho_w\frac{\mathbf{k}k_{rw}}{\mu_w}\left(\nabla P - \rho_w g\nabla Z\right)\right] + Q_w = 0 \tag{4.2a}$$

$$\frac{\partial}{\partial t}\left[\varphi\rho_o S_o\right] - \nabla\cdot\left[\rho_o\frac{\mathbf{k}k_{ro}}{\mu_o}\left(\nabla P - \rho_o g\nabla Z\right)\right] + Q_o = 0 \tag{4.2b}$$

in which $\mathbf{k}$ is the permeability tensor, k_{rj} is the relative permeability of phase j, μ_j is the viscosity of phase j ' and Z is the depth of the reservoir (downwards positive). g is the gravitational acceleration. Q_w and Q_o represent sources/sinks due to wells. Darcy's Law is used to represent phase velocities in (4.2).

To fully specify the flow conditions the boundary conditions must be specified. As mentioned above no-flow conditions prevail in this work:

$$\mathbf{u}_j\cdot\mathbf{n} = 0, \qquad j\in\{w,o\}, \tag{4.3}$$

in which $\mathbf{u}_j$ is the velocity of phase j and $\mathbf{n}$ is an outward pointing normal vector.

The phase velocities are computed by Darcy's law. The flow potential Φ and Darcy velocities are defined as:

$$\nabla\Phi_j = \nabla P - \rho_j g\nabla Z \tag{4.4}$$

$$\mathbf{u}_j = -\mathbf{k}\lambda_j\nabla\Phi_j, \tag{4.5}$$

in which $\lambda_j = \frac{k_{rj}}{\mu_j}$ is the relative phase mobility. The permeability tensor is assumed diagonal. The phase mass fluxes driven by pressure and gravitational forces are introduced as:

$$\mathbf{q}_j^F = -\rho_j\mathbf{k}\lambda_j\nabla P \tag{4.6}$$

$$\mathbf{q}_j^G = \rho_j^2\mathbf{k}\lambda_j g\nabla Z \tag{4.7}$$

4.1.1 Well models

Until now only the external boundaries of the reservoir have been discussed. However, wells constitute another type of boundary and must be dealt with properly. The measurable production data for history matching are measured at the wells or after the fluids have left the wells. It is evident that the treatment of wells plays a central role in history matching and in reservoir simulation in

general. The scale of the grid blocks used in reservoir simulation is often orders of magnitude larger than the scale of wells. Thus, a $50m \times 50m \times 10m$ grid block may for instance contain a well with radius $15cm$. The numerical flow simulation cannot resolve the small scale features related to the well which makes it less apparent how to deal with wells in reservoir simulators. One solution may be to refine the grid around wells. This can be achieved relatively easily by the use of unstructured grids, e.g. Voronoi grids. Traditionally, this approach has not been applied to a large extent since reservoir simulation in the past, and to some extend also today, has required the use of structured grids. Instead, a well model ensures that the numerically computed well pressure is identical to an analytically derived pressure. The analytically derived pressure is exact for radial flow of a single phase near the well (Aziz & Settari 1979).

With the convention of positive mass flow out of the reservoir, the general mass flux terms for the wells are defined by:

$$\text{Phase } j \text{ from well } \omega: \qquad Q_j^\omega = WI^\omega \rho_j \lambda_j \left(P - P^\omega\right) \tag{4.8}$$

$$\text{Net from well } \omega: \qquad Q^\omega = Q_w^\omega + Q_o^\omega, \tag{4.9}$$

in which P^ω denotes the pressure in the well. Equation 4.8 is referred to as the well equation.

Agreement with the analytically calculated pressure is enforced through the well index, WI. For a vertical well the analytically derived pressure is given as:

$$P^\omega = P_o + \frac{\mu Q^\omega / \rho}{\theta H (k_x k_y)^{1/2}} \ln\left(\frac{r_w}{r_o}\right), \tag{4.10}$$

where r_o denotes the radial position at which the pressure is equal to the numerically calculated block pressure. P_o is the pressure at r_o, H is the height of the well, k_x and k_y are permeability in the x-direction and the y-direction, respectively. r_w denotes the radius of the well. For a non-square block-centered cartesian grid block Peaceman (1983) gives the following expression to determine r_o:

$$r_o = 0.28 \frac{\left[\left(\frac{k_y}{k_x}\right)^{1/2} \Delta x^2 + \left(\frac{k_x}{k_y}\right)^{1/2} \Delta y^2\right]}{\left(\frac{k_y}{k_x}\right)^{1/4} + \left(\frac{k_x}{k_y}\right)^{1/4}} \tag{4.11}$$

For this model the well index in (4.8) is given by:

$$WI = \left[\frac{\theta (k_x k_y)^{1/2} H}{\ln\left(\frac{r_o}{r_w}\right) + s}\right], \tag{4.12}$$

where s is a skin factor used to account for damage to the formation in the vicinity of the well.

This methodology can be generalized to wells covering multiple grid blocks. In this case a well index has to be defined for each grid block containing the well. The term multiple-block well completion denotes wells covering multiple grid blocks. The pressure in a well is needed when well rates are computed. If a well covers multiple grid blocks it is necessary to include the pressure of the fluid column when the pressure in a well segment is calculated.

In this work the following assumptions regarding the flow in the wells have been made:

- No friction between liquids and well surface
- Wells are either vertical or horizontal - sloping wells are not an option
- Variable pressure gradient in the well
  - Pressures in well segments are calculated from liquid inflows in each overlying well segment - see equation (4.13).

The first assumption ensures that there is no pressure drop through the well caused by friction. This simplifies the handling of the wells since special treatment with advanced pipe flow models is avoided. In the current implementation of the BOSS simulator sloping wells cannot be handled. However, the use of vertical and horizontal wells is sufficient for the applications in this dissertation. The pressure exerted by overlaying well segments is calculated weighting the fluid densities with the inflow rates for the concerned fluid. The well pressure in the N'th segment is given by:

$$P_N = P_1 + \frac{g}{2}\sum_{i=1}^{N-1}\left(\frac{\lambda_{(w,i)}\rho_{(w,i)} + \lambda_{(o,i)}\rho_{(o,i)}}{\lambda_{(w,i)} + \lambda_{(o,i)}}\right)(z_{i+1} - z_i), \tag{4.13}$$

where P_1 is the pressure in the top grid block, $\lambda_{(p,i)}$ is the relative phase mobility of phase p in the grid block containing the i'th well segment. Figure 4.1 shows a schematic picture of a vertical multiple-block well completion.

The flow rate of phase p in well segment i is then given as:

$$q_p^{(w,i)} = (WI)_i\lambda_p(P_p - P_p^{(w,i)}), \quad WI = \frac{2\pi kh}{\ln\frac{r_o}{r_w} + s} \tag{4.14}$$

4.2 Discretization

This section serves the purpose of introducing the very basics of the numerical method used to solve the flow equations. A basic knowledge about the numerics involved with reservoir simulation is needed in order to understand the background of the adjoint based sensitivity calculation discussed in section 4.3 which plays a cental role in history matching applications discussed later.

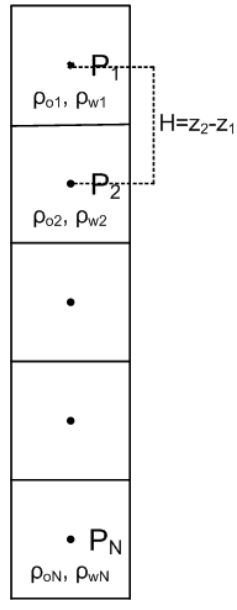


Figure 4.1: Schematic of a vertical well covering multiple blocks.

The governing flow equations (4.2) with the boundary conditions (4.3) are solved using a finite volume representation of the discretized equations. The grid is a variable volume block centered grid. For an elaborate discussion on this topic the reader may consult Johansen & Kristensen (2006) or a textbook on numerical methods for reservoir simulation, e.g. Aziz & Settari (1979).

Spatial and temporal discretization of the flow equations yields (Aziz, Durlofsky & Tchelepi 2005):

$$\boldsymbol{T}^{n+1}\boldsymbol{u}^{n+1} - \boldsymbol{D}(\boldsymbol{u}^{n+1} - \boldsymbol{u}^{n}) - \boldsymbol{G} - \boldsymbol{Q} = 0 = \boldsymbol{R}, \tag{4.15}$$

where $\boldsymbol{u} = [p_1 \ S_1 \ ... \ p_N \ S_N]$ is the vector of unknowns which in the fully implicit formulation consists of grid block pressures and (water) saturations. $\boldsymbol{T}$ is the matrix of transmissibilities, $\boldsymbol{D}$ contains accumulation terms, $\boldsymbol{G}$ contains gravity terms, and $\boldsymbol{Q}$ represents sink/source terms. $\boldsymbol{R}$ is the residual which is driven to zero by a non-linear solver.

In the fully implicit formulation the accumulation term can be expressed as:

$$\boldsymbol{D}(\boldsymbol{u}^{n+1} - \boldsymbol{u}^{n}) = \frac{\boldsymbol{V}}{\Delta t}(\boldsymbol{\phi\rho S}|_{t+\Delta t} - \boldsymbol{\phi\rho S}|_{t}) \tag{4.16}$$

The matrix $\boldsymbol{T}$ consists of seven diagonals with each element being a 2×2 matrix for a water-oil system in three dimensions. The residual is driven to zero

by a Newton-type nonlinear solver which necessitates the computation of the Jacobian of the residual:

$$\boldsymbol{J} = \frac{\partial \boldsymbol{R}}{\partial \boldsymbol{u}} = \begin{pmatrix} \frac{\partial r_1}{\partial p_1} & \frac{\partial r_1}{\partial S_1} & \frac{\partial r_1}{\partial p_1} & \cdots & \frac{\partial r_1}{\partial S_{N_{blk}}} \\ \frac{\partial r_2}{\partial p_1} & \frac{\partial r_2}{\partial S_1} & \frac{\partial r_2}{\partial p_1} & \cdots & \frac{\partial r_2}{\partial S_{N_{blk}}} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial r_{2N_{blk}}}{\partial p_1} & \frac{\partial r_{2N_{blk}}}{\partial S_1} & \frac{\partial r_{2N_{blk}}}{\partial p_1} & \cdots & \frac{\partial r_{2N_{blk}}}{\partial S_{N_{blk}}} \end{pmatrix} \tag{4.17}$$

The Jacobian has the same structure as the transmissibility matrix and it is therefore stored in a sparse format. The Jacobian plays an important role in the computation of sensitivities for history matching as will be demonstrated later in section 4.3.

In order to clarify the adjoint approach described in section 4.3 and in appendix A we will now consider how the flux between two grid blocks is approximated. For simplicity it is assumed that the flow is taking place in the x-direction. Generalization to other directions is straightforward. Figure 4.2 shows a schematic of a one-dimensional grid. The distances between the center of grid block i, j to the centers of the neighboring grid blocks are denoted Δx_+ and Δx_-.

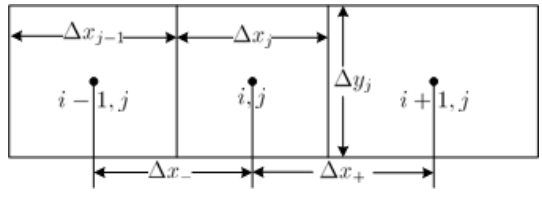


Figure 4.2: Schematic of a one-dimensional grid.

The flux over the face between grid blocks (i, j) and $(i + 1, j)$ is given by:

$$q^F_{i+1/2,j} = H_{res}\Delta y_j [k_x]_{i+1/2,j} [\rho\lambda]_{i+1/2,j} \frac{p_{i+1,j} - p_{i,j}}{\Delta x_{i+}} = \Upsilon_{i+1/2,j}(p_{i+1,j} - p_{i,j}), \tag{4.18}$$

where $\Upsilon_{i+1/2,j}$ denotes the interface transmissibility. H_{res} is the reservoir height. for a three-dimensional grid H_{red} is replaced with Δz_j. The trans-

missibility can be divided into a geometric part and a fluid part:

$$\Upsilon_{i+1/2,j} = \Gamma_{i+1/2,j} H_{i+1/2,j} \tag{4.19a}$$

$$\Gamma_{i+1/2,j} = \frac{H_{res} \Delta y_j k_{x,i+1/2,j}}{\Delta x_{i+}} \tag{4.19b}$$

$$H_{i+1/2,j} = \left(\frac{\rho k_r}{\mu} \right)_{i+1/2,j} \tag{4.19c}$$

The geometric transmissibility, Γ , depends on the type of descretization and the physical properties of the reservoir rock and is a static property which can be calculated in a pre-processing step prior to simulation. The fluid transmissibility (H), varies with pressure and saturation.

The interface permeability in (4.19b) is computed from harmonic averages of the grid-block permeabilities:

$$k_{x,i+1/2,j} = \frac{\Delta x_i + \Delta x_{i+1}}{\dfrac{\Delta x_i}{k_{x,i,j}} + \dfrac{\Delta x_{i+1}}{k_{x,i+1,j}}} \tag{4.20a}$$

The use of a harmonic average to compute interface permeabilities enforces flux continuity across the interface. The fluid transmissibility, H, is upwinded to the direction of flow as this approach has preferable stability properties (Aziz & Settari 1979).

4.3 Adjoint sensitivity calculation

In chapters 7 and 8 the use of sensitivities of production data with respect to permeability and porosity plays a central role. Calculation of sensitivities is done using an adjoint approach which allows for an efficient calculation of the sensitivities. This section introduces the method and presents its mathematical background along with a few example calculations. A detailed description of the computation of the derivatives needed for the adjoint approach is provided in Appendix A.

The sensitivity of a production data, β, with respect to the property $\boldsymbol{m}$ is defined as:

$$\boldsymbol{s} = \frac{\partial \beta}{\partial \boldsymbol{m}}. \tag{4.21}$$

$\boldsymbol{m}$ is marked as a vector since the property in question is typically defined at the grid block scale. The computation of sensitivities of production data can be computationally heavy. However, the introduction of the so-called *adjoint* approach has reduced the computational load significantly compared to previous

methods for calculation of sensitivity coefficients (Li et al. 2003). The cost of computing sensitivities is now comparable to that of doing a reservoir simulation (P. Sarma 2005).

We are now setting up the basis to calculate the sensitivity of the scalar function β with respect to a reservoir property $\boldsymbol{m}$, e.g. absolute permeabilities or porosities at the grid block scale. The function β will be treated as a function of the vector of unknowns, $\boldsymbol{u}$ and $\boldsymbol{m}$, i.e.

$$\beta = \beta(\boldsymbol{u}^1, ..., \boldsymbol{u}^L, \boldsymbol{m}), \tag{4.22}$$

where L is the last time step where the sensitivity is desired, which is usually at the end of the simulation period. The adjoint formulation is based on the formation of an adjoined system of equations in which the discretized model equations (4.15) are acting as constraints. Adjoining the discretized equations to β yields:

$$J = \beta + \sum_{n=0}^{L} (\boldsymbol{\lambda}^{n+1})^T \boldsymbol{R}^{n+1}, \tag{4.23}$$

where $\boldsymbol{\lambda}^{n+1}$ is the vector of adjoint variables, or alternatively, a vector of Lagrange multipliers for the model constraint at time step $n+1$. The adjoint variables should not be confused with the mobility which shares the same symbol. It should be noted, however, that the adjoint variables are typed in bold because they are vectors. $\boldsymbol{R}$ is the residual from the discretized flow equations.

Since the residual is driven to zero by the nonlinear solver in the simulator the following equation holds (Wu et al. 1998):

$$\nabla_{\boldsymbol{m}} J = \nabla_{\boldsymbol{m}} \beta. \tag{4.24}$$

The size of the adjoint vector is similar to the size of the vector of unknowns, i.e. it has the length $2N + N_{inj}$ where N is the number of grid blocks and N_{inj} is the number of injectors. Here it is assumed that producers are pressure constrained and that a fully implicit formation of the discretized flow equations is used. In the following derivation the notation used in (Wu et al. 1998) and (Li et al. 2003) is applied.

The total differential of the adjoint system of equations (4.23) is calculated:

$$dJ = d\beta + \sum_{n=0}^{L} \left[(\boldsymbol{\lambda}^{n+1})^T (\nabla_{\boldsymbol{u}^{n+1}} (\boldsymbol{R}^{n+1})^T)^T d\boldsymbol{u}^{n+1} + (\nabla_{\boldsymbol{m}} (\boldsymbol{R}^{n+1})^T)^T d\boldsymbol{m} \right] + \sum_{n=0}^{L} (\boldsymbol{\lambda}^{n+1})^T (\nabla_{\boldsymbol{u}^n} (\boldsymbol{R}^{n+1})^T)^T d\boldsymbol{u}^n. \tag{4.25}$$

Changing the sum to start from index 1 gives:

$$dJ = d\beta + \sum_{n=1}^{L} \big[((\boldsymbol{\lambda}^n)^T(\nabla_{\boldsymbol{u}^n}(\boldsymbol{R}^n)^T)^T + (\boldsymbol{\lambda}^{n+1})^T(\nabla_{\boldsymbol{u}^n}(\boldsymbol{R}^{n+1})^T)^T)d\boldsymbol{u}^n + (\boldsymbol{\lambda}^n)^T(\nabla_{\boldsymbol{m}}(\boldsymbol{R}^n)^T)^T d\boldsymbol{m}\big] + BT, \tag{4.26}$$

where

$$BT = (\boldsymbol{\lambda}^{n+1})^T[(\nabla_{\boldsymbol{u}^{L+1}}(\boldsymbol{R}^{L+1})^T)^T d\boldsymbol{u}^{L+1} + (\nabla_{\boldsymbol{m}}(\boldsymbol{R}^{n+1})^T)^T d\boldsymbol{m}] + (\boldsymbol{\lambda}^1)^T(\nabla_{\boldsymbol{u}^0}(\boldsymbol{R}^1)^T)^T d\boldsymbol{u}^0. \tag{4.27}$$

The initial conditions are invariant which implies that $d\boldsymbol{u}^0 = 0$. If $\boldsymbol{\lambda}^{L+1}$ is chosen to be 0 the BT term is obviously zero. This will serve as the end condition for the adjoint problem.

Lets us now take the total differential of the scalar parameter β:

$$d\beta = \sum_{n=1}^{L} \big[(\nabla_{\boldsymbol{u}^n}\beta)^T d\boldsymbol{u}^n\big] + (\nabla_{\boldsymbol{m}}\beta)^T d\boldsymbol{m}. \tag{4.28}$$

Equation (4.28) is combined with equation (4.26):

$$dJ = \sum_{n=1}^{L} \big[(\nabla_{\boldsymbol{u}^n}\beta)^T + (\boldsymbol{\lambda}^n)^T(\nabla_{\boldsymbol{u}^n}(\boldsymbol{R}^n)^T)^T + (\boldsymbol{\lambda}^{n+1})^T(\nabla_{\boldsymbol{u}^n}(\boldsymbol{R}^{n+1})^T)^T d\boldsymbol{u}^n + (\boldsymbol{\lambda}^n)^T(\nabla_{\boldsymbol{m}}(\boldsymbol{R}^n)^T)^T d\boldsymbol{m}\big] + (\nabla_{\boldsymbol{m}}\beta)^T d\boldsymbol{m}. \tag{4.29}$$

The essential operation in the adjoint approach is to simplify the total differential in equation (4.29) by forcing the terms multiplying the $d\boldsymbol{u}$ terms to vanish. This is done by equating the proper terms to zero, i.e.:

$$(\nabla_{\boldsymbol{u}^n}\beta)^T + (\boldsymbol{\lambda}^n)^T(\nabla_{\boldsymbol{u}^n}(\boldsymbol{R}^n)^T)^T + (\boldsymbol{\lambda}^{n+1})^T(\nabla_{\boldsymbol{u}^n}(\boldsymbol{R}^{n+1})^T)^T = 0. \tag{4.30}$$

Transposing equation (4.30) gives:

$$\boxed{(\nabla_{\boldsymbol{u}^n}(\boldsymbol{R}^n)^T)\boldsymbol{\lambda}^n = -(\nabla_{\boldsymbol{u}^n}(\boldsymbol{R}^{n+1})^T)\boldsymbol{\lambda}^{n+1} - (\nabla_{\boldsymbol{u}^n}\beta)^T,} \tag{4.31}$$

which is the system of adjoint equations. If λ^n is consistent with the adjoint equation (4.31) the sensitivity of β with respect to the parameter vector $\boldsymbol{m}$ can be calculated from equation (4.29) which reduces to:

$$dJ = \sum_{n=1}^{L} \big[(\boldsymbol{\lambda}^n)^T(\nabla_{\boldsymbol{m}}(\boldsymbol{R}^n)^T)^T d\boldsymbol{m}\big] + (\nabla_{\boldsymbol{m}}\beta)^T d\boldsymbol{m}. \tag{4.32}$$

The sensitivity is therefore given as (remember equation (4.24)):

$$\boxed{\frac{dJ}{d\boldsymbol{m}} = \nabla_{\boldsymbol{m}}\beta = \sum_{n=1}^{L}\left[(\boldsymbol{\lambda}^n)^T(\nabla_{\boldsymbol{m}}(\boldsymbol{R}^n)^T)\right] + (\nabla_{\boldsymbol{m}}\beta).} \tag{4.33}$$

Let us dwell a little on the structure of the adjoint system (4.31). The equation is linear in the Lagrange multipliers (or equivalently the adjoint variables). The multipliers may be determined by stepping backwards in time starting from $t = t_L$. The condition $\boldsymbol{\lambda}^{L+1} = 0$ is used to initialize the system. The matrix $(\nabla_{\boldsymbol{u}^n}(\boldsymbol{R}^n)^T)$ multiplying the lagrange multiplier $\boldsymbol{\lambda}^n$ can be identified as the transposed Jacobian of the discretized model equations. The Jacobian is formed during the forward simulation to be used in the nonlinear solver (which is the Newton method) and is therefore known. The other matrix occurring in equation (4.31) multiplying $\boldsymbol{\lambda}^{n+1}$ is a diagonal matrix related to the accumulation term in the model equations. This is due to the implicit nature of the numerical method. The only part in equations (4.15) and (4.16) that depend on $\boldsymbol{u}^n$, i.e. the current vector of unknowns, is the accumulation term. All other terms are evaluated implicitly in the next time step, i.e. as functions of $\boldsymbol{u}^{n+1}$.

The scalar function β is chosen by the user and the choice of β depends on the specific problem at hand. If the adjoint method is used in a minimization problem β might be chosen as the minimized function, i.e. the objective function. In general reservoir engineering applications β may be the bottom hole pressure at injectors or the watercut. If one needs the calculation of several sensitivities, e.g. for bottom hole pressures for multiple injectors, an adjoint equation system has to be formulated for each pressure. However, it is important to recognize that only the right-hand sides of equation (4.31) change if several systems are included. This means that only one factorization of the left-hand side matrix is required which reduces the computational work significantly. If β does not depend explicitly on $\boldsymbol{m}$ the derivative $\nabla_{\boldsymbol{m}}\beta$ in the sensitivity equation (4.33) will disappear.

The details of the adjoint sensitivity calculation are given in Appendix A.

The current implementation of the adjoint procedure supports the following features:

- Sensitivity of watercut, oil rate, and injector bottomhole pressure with respect to:
  - Permeability
  - Porosity
- Three-dimensional block centered grids
- Vertical and horizontal well completions

4.3.1 Examples of sensitivity calculations

A few sample calculations of sensitivity with respect to grid block permeability and porosity are presented in this section. In general, sensitivities can be difficult to interpret when the reservoir is heterogeneous and when multiple wells are involved. Hence, only relatively simple setups are considered here in order to accommodate the clarity of the examples. The first case presented here is a simple 1D setup with ten grid blocks. Water is being injected in grid block number 1 and production is taking place in block number 10. Oil and water are treated as incompressible. Figure 4.3 shows calculated sensitivities of the injector pressure with respect to absolute permeability at the beginning of the simulation. Numerical results obtained from perturbations (finite differences) are indicated by the symbol $\times$. The sensitivities calculated with the adjoint method agree with the numerical results. In this case the difference is in the order of 10^{-2}. The figure indicates that the grid blocks close to the grid block containing the injector are most influential on the bottom hole pressure. The sensitivity associated with the injector grid block is smaller than the sensitivities associated with the grid block lying between the two wells. This is a consequence of the block centered gridding used in the BOSS simulator. A well is placed in the middle of a block which means that the distance the injected fluid can travel inside a well-containing block is less than half of the grid block length - the exact length depends on the well diameter but is always smaller than half of the grid block length. This explains the counter-intuitive observation that the blocks containing wells are less sensitive to changes in permeability than the grid blocks in between.

The second example is based on a two-dimensional reservoir with a quarter-nine-spot well configuration. The specifications of the wells are given in Table 4.1. The reservoir has dimensions $100m \times 100m \times 10m$ and permeability is $50mDa$ over the entire domain. Porosity is 0.25. The flow is solved on a $50 \times 50 \times 1$ grid. Water is being injected at a constant rate in three corners and production is taking place in the last corner. Oil and water are treated as incompressible and the reservoir is fully oil saturated initially. Figure 4.4 shows the calculated sensitivity of bottomhole pressure at the three injectors with respect to grid block permeability. The shown sensitivities are scaled in order to enhance the visual quality. It is evident that the permeabilities in the grid blocks close to the wells are the most influential. Because of symmetry in the well configuration the sensitivity of pressure in injectors 2 and 3 are symmetric around the diagonal going from injector 1 to the producer.

As a last example of a sensitivity calculation a simple three-dimensional problem is considered. A symmetry element of a five-spot pattern is applied for the well configuration. This involves only one injector and one producer. The specifications of the wells are given in Table 4.2. The wells are completed in the z-direction and cover all three layers. The reservoir has dimensions $100m \times 100m \times 10m$ and permeability is $50mDa$ over the entire domain. Porosity is

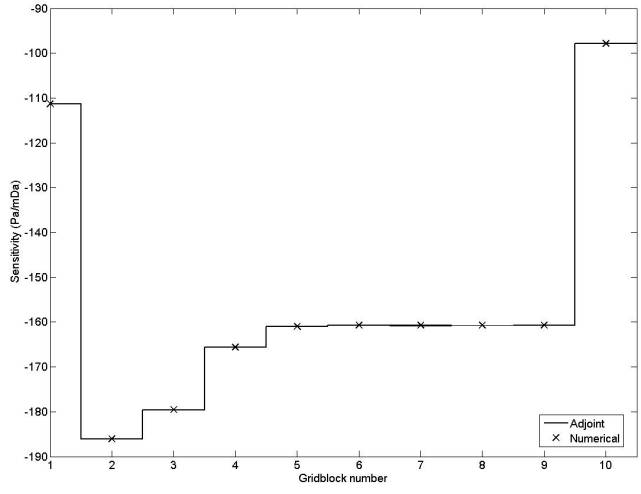


Figure 4.3: Sensitivity in the 1D case of injector BHP with respect to grid block permeability. $\times$ indicates sensitivities obtained from finite difference approximations.

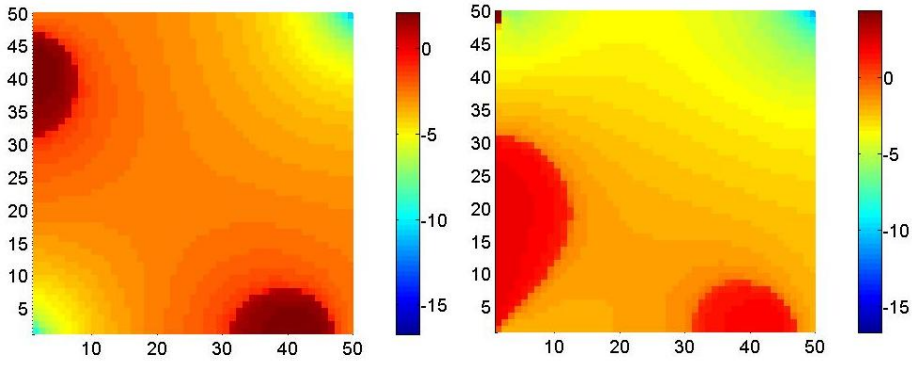
Table 4.1: Specifications of the wells. The wells are placed in a quarter-nine-spot pattern. Only water is being injected.

Well	Type	i	j	Constraint
I1	Inj	1	50	Rate ($5m^3/day$)
I2	Inj	1	50	Rate ($5m^3/day$)
I3	Inj	50	1	Rate ($5m^3/day$)
P1	Prod	50	50	BHP ($10bar$)

0.25. The flow is solved on a $10 \times 10 \times 3$ grid. Figure 4.5 shows the calculated sensitivity of oil rate with respect to grid block permeability and porosity. The sensitivities are symmetric around the diagonal connecting the wells which is expected because of the symmetric well configuration and the homogeneity of the field.

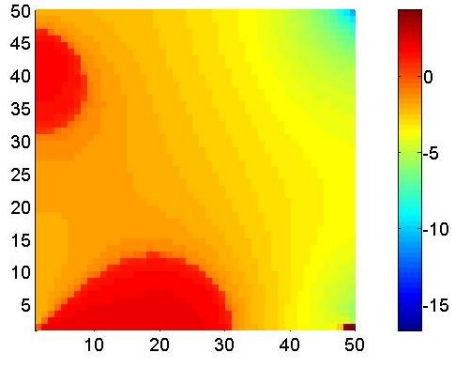
Table 4.2: Specifications of the wells for the three-dimensional setup. The wells are placed in a symmetry element of a 5-spot pattern. Only water is being injected.

Well	Type	i	j	k_1	k_2	Constraint
I1	Inj	1	50	1	3	Rate ($5m^3/day$)
P1	Prod	50	50	1	3	BHP ($10bar$)



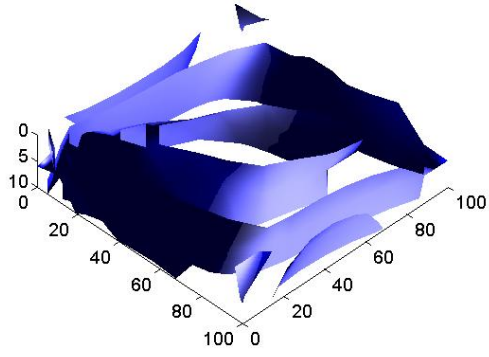
(a) Injector 1.

(b) Injector 2.

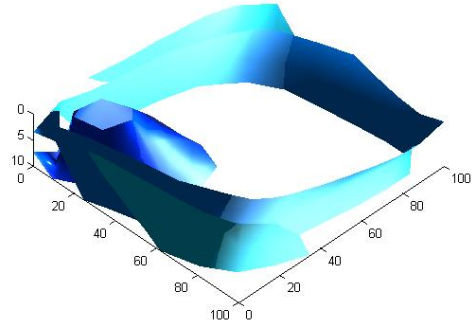


(c) Injector 3.

Figure 4.4: Sensitivity of bottomhole pressures in the three injectors with respect to grid block permeability. A scaling has been applied to the sensitivities in order to improve the visual quality.



(a) W.r. to permeability.



(b) W.r. to porosity.

Figure 4.5: Sensitivity of oil rate with respect to grid block permeability and porosity.

Part II

Methods for history matching of oil production

Chapter 5

Gradual deformation

The gradual deformation method was briefly introduced in the literature review in Chapter 2. A more detailed description is provided in the present chapter and application of the gradual deformation method to synthetic history matching examples is demonstrated.

5.1 Description of the method

Gradual deformation was introduced by Hu (2000) as a technique to constrain history matches to simple statistics. In Hu (2002) and Hu & le Ravalec-Dupin (2004) the method is further developed. The basics of the method is presented in the following sections. For a comprehensive description on more advanced uses of gradual deformation the reader may consult the papers from Hu and his coworkers mentioned above.

Consider a realization of a spatially distributed property denoted by Z_i to imply that it is the ith realization from a stochastic process. In general, a vast number of realizations which all honor given second order statistics (variogram) exist. Therefore we need to distinguish between the individual realizations. The essential operation in gradual deformation is to form a new realization from a set of old ones as a simple linear combination:

$$\boldsymbol{Z}_{new} = \sum_{i=1}^{N_r} \alpha_i Z_i, \tag{5.1}$$

where α_i is a weighting factor. (Hu 2000) noticed that certain constraints on

the geology of the resulting realization, Z_{new} could be realized by imposing simple constraints on the weighting factors and on the realizations Z_i. If the realizations Z_i have the same variogram then the variogram will be conserved in the new realization, if:

1. The old realizations, if needed, are transformed to being normally distributed and having zero mean
2. $\sum_{i=1}^{N_r} \alpha^2 = 1$

If only two realizations are used to form the new realization it is trivial to see that the expression

$$\boldsymbol{Z}_{new} = \sin(\pi\rho)\boldsymbol{Z}_1 + \cos(\pi\rho)\boldsymbol{Z}_2 \tag{5.2}$$

fulfills the requirements that the sum of squared weights is unity. In general, if the number of old realizations is N_r the weights may be found as an N-dimensional hypersphere (Hu 2000):

$$\alpha_1 = \prod_{i=1}^{n-1} \cos\rho_i\pi \tag{5.3a}$$

$$\alpha_{i+1} = \sin\rho_i\pi \prod_{j=i+1}^{n-1} \cos\rho_j\pi, \quad i = 1, ..., n-2 \tag{5.3b}$$

$$\alpha_n = \sin\rho_{n-1}\pi. \tag{5.3c}$$

Equation (5.2) illustrates that the parameter ρ can be used to deform the resulting realization. If $\rho = 0$ the resulting realization will be equal to $\boldsymbol{Z}_2$ and for $\rho = \frac{1}{2}$ $\boldsymbol{Z}_1$ will be reproduced. For $\rho \in]0, \frac{1}{2}[$ and $\rho \in]\frac{1}{2}, 1[$ the result will be some mixture of the two old realizations. Consequently, ρ is referred to as the *deformation parameter*. Figure 5.1 shows a gradual deformation of a spatially distributed property. Two realizations are used in the gradual deformation. Both are generated by sequential Gaussian simulation and are constrained to the same variogram. The chosen value of the deformation parameter results in a realization which is a mix between the two old realizations. The two old realizations both exhibit a trend in the North-East direction which is reproduced in the gradually deformed result. The gradually deformed result conserves the second order statistics of the old fields. Consequently, the variogram is conserved and the result is equiprobable to the old realizations with respect to second order statistics.

As illustrated above the gradual deformation techniques is a convenient way of generating equiprobable realizations of Gaussian distributions. However, the main advantage of the method is that it can be used to reformulate history matching. History matching is essentially an optimization problem where the

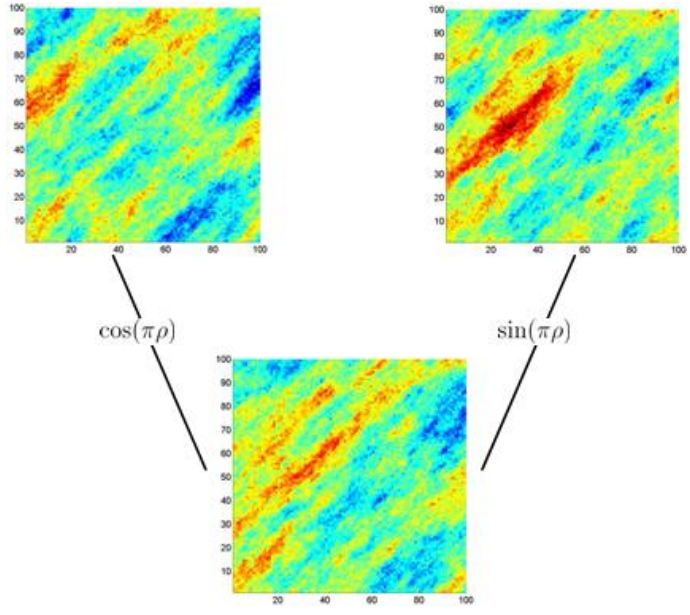


Figure 5.1: Gradual deformation made with two realizations of permeability. The deformation parameter, ρ is 0.25. The realizations are made with sequential Gaussian simulation in the geostatistical software package GSLIB (Deutsch & Journel 1998).

production mismatch is to be minimized by adjusting certain physical properties of the reservoir and reservoir fluids. If permeability is the only adjustable parameter the problem of history matching can be stated by the general formulation:

$$\mathbf{k}^* = \underset{\mathbf{k}}{\text{Argmin}} \left[E(\mathbf{k}) \right], \tag{5.4}$$

where E is the objective function quantifying the production data mismatch.

In cases where permeability can be represented by a Gaussian distribution it is possible to parameterize the permeability field by applying gradual deformation. Thus, permeability can be fully described by a set of equiprobable realizations and a set of deformation parameters. If the simple expression in equation (5.2) is used for the deformation the permeability can be represented by two realizations and only one deformation parameter. Thus, permeability can be represented by an operator performing a gradual deformation. The gradual deformation is a function of a number of equiprobable realizations and the deformation parameters:

$$\mathbf{k} = \mathcal{G}(\boldsymbol{Z}_i, \rho_j), \quad i = 1, 2..., N_{rel}, \quad j = 1, 2, ..., N_{rel} - 1. \tag{5.5}$$

The realizations $\boldsymbol{Z}_i$ are generated by a geostatistical simulator and are fixed. This leaves the deformation parameters as the only adjustable parameters. The idea behind gradual deformation in history matching is to search for the optimal deformation parameters which minimizes the production data mismatch:

$$\rho^* = \underset{\rho}{\text{Argmin}} \left[E(\rho) \right], \tag{5.6}$$

In the simplest case where only two realizations are used in the gradual deformation the number of deformation parameters is only one. This reduces the history matching to a one-dimensional optimization problem. By parameterizing the permeability with the gradual deformation method the number of independent variables is reduced significantly. Simulation models may have hundreds of thousands of grid blocks which leaves the same number of independent variables in the optimization problem. If additional parameters are adjusted the number of independent variables is increased further.

The constraints on the deformation parameters ensure that the second order statistics are conserved in gradual deformation. However, the definitions in equations (5.1) and (5.2) do not conserve hard data. Therefore, conditioning to hard data has to be done in a separate step. In Hu (2002) the gradual deformation method is further developed such that the method also works with dependent realizations. Apart from requiring that the deformation parameters squared sum to unity the parameters also need to sum to unity in the power 1 in order to conserve hard data. Thus, the following constraints are imposed on the deformation parameters:

$$\sum \alpha^2 = 1 \tag{5.7a}$$

$$\sum \alpha = 1. \tag{5.7b}$$

The additional constraint takes out one degree of freedom. Consequently, the simple expression in equation (5.2) only holds for $\rho = 0$ and $\rho = 1$. In order to increase the degrees of freedom the number of realizations used to perform gradual deformation has to be increased. Hu (2002) suggests the following expressions for gradual deformation of dependent realizations:

$$\alpha_1 = \frac{1}{3} + \frac{2}{3} \cos \rho\pi \tag{5.8a}$$

$$\alpha_2 = \frac{1}{3} + \frac{2}{3} \sin(\pi(-\frac{1}{6} + \rho)) \tag{5.8b}$$

$$\alpha_3 = \frac{1}{3} + \frac{2}{3} \sin(\pi(-\frac{1}{6} - \rho)), \tag{5.8c}$$

and the new realization is given as

$$\boldsymbol{Z} = \alpha_1 \boldsymbol{Z}_1 + \alpha_2 \boldsymbol{Z}_2 + \alpha_3 \boldsymbol{Z}_3. \tag{5.9}$$

This formulation involves three old realizations but still only one deformation parameter and ensures that hard data is conserved along with the second order statistics without a separate conditioning step. According to Hu (2002) gradual deformation of dependent realizations compared to gradual deformation of independent realization ensures better numerical stability of the method.

5.1.1 Gradual deformation in history matching

As described above, the gradual deformation technique can be used to parameterize the optimization problem related to history matching. Apart from simplifying the optimization problem this parameterization makes it easy to constrain the history match to hard data and variogram. An important detail must be addressed before a history matching procedure is presented. In order to fulfill the requirements for gradual deformation the deformed property has to have zero mean and a standard normal distribution. This is not the case for permeability and porosity which cannot take negative values. Consequently, deformation has to take place on transformed variables. After deformation the result can be backtransformed to physically meaningful values.

Hu (2000) and Hu (2002) suggest a history matching algorithm of the following form:

1. Generate initial permeability field, $\boldsymbol{Z}_1$ to initialize history matching
2. Initiate $\rho = 0$
3. set $i = 0$
4. Compute permeability fields $\boldsymbol{Z}_2$ and $\boldsymbol{Z}_3$
5. Set $k = 0$
   (a) Perform gradual deformation with the three realizations, $\boldsymbol{Z}_1$, $\boldsymbol{Z}_2$, $\boldsymbol{Z}_3$ and ρ
   (b) Backtransform the gradually deformed field to natural values
   (c) Perform fluid flow simulation on the backtransformed field
   (d) Update ρ according to the production data mismatch
   (e) $k = k + 1$
   (f) If $k = k_{max}$ or other stop criteria is met, stop inner loop, else go to (a)
6. If the objective function decreased set $\boldsymbol{Z}_1$ equal to the permeability field associated with the best match
7. $i = i + 1$

8. If $i = i_{max}$ or other stop criteria is met, stop, else go to 4

The above algorithm consists of two loops, an outer loop where the realizations of permeability $\boldsymbol{Z}_2$ and $\boldsymbol{Z}_3$ are changed and an inner loop where the optimal deformation parameter is searched for. Computation of two new realizations is done by changing the seeds used to initialize the sequential simulation. The realization which gives the best match of the production data is stored in the field $\boldsymbol{Z}_1$. Thus, the best realization is carried forward to the next outer loop in $\boldsymbol{Z}_1$ and it is reproduced when $\rho = 0$. Termination of the inner loop is typically done when a maximum number of inner iterations has been reached. The outer loop is terminated when the production mismatch has reached a specified tolerance or when a maximum number of outer iterations is reached. Hu (2000) uses a gradient based optimization method in the inner loop to search for an optimal deformation. Such an approach requires that the objective function is continuous with respect to changes in the model parameters. This may be the case for smoothly varying Gaussian fields but for facies models or truncated Gaussian models the objective function is often discontinuous. In such cases a derivative-free optimizer like bisection or the Dekker-Brent method is needed. The implementation of gradual deformation used in this dissertation applies the one-dimensional, derivative-free Dekker-Brent method in the inner optimization loop. In appendix B the method is presented in detail. The implemented gradual deformation procedure makes use of the sequential Gaussian simulator found in geostatistical software library GSLIB (Deutsch & Journel 1998). Sequential simulation is done using transformed variables witch are normally distributed with zero mean. The fact that normal scores are applied in sequential Gaussian simulation is exploited in the history matching procedure since gradual deformation also requires that the variables used for deformation are normally distributed. Backtransformation is only done when a reservoir simulation is performed.

In the following section the gradual deformation method is applied to simple synthetic history matching cases.

5.2 Application of gradual deformation in history matching cases

The first case involves history matching of a heterogeneous field by the use of the gradual deformation method. The field considered here is a small symmetry element of quarter-nine-spot well configuration. The reservoir is quadratic and a gridding of 10×10 is used with injectors in three corners and a producer in the last corner. Permeability is the only unknown property and porosity is assumed constant in the entire domain. Watercut in the producer is matched together with injector bottomhole pressures for the three injectors. Production data is

measured at eight times. Specifications of the well constraints are provided in Table 5.1.

One important assumption is made regarding the statistics of the heterogeneous permeability fields: perfect knowledge about the statistics of permeability is assumed. This is important since transformation of the permeability is needed when the gradual deformation is performed because the fields used in the method have to be normally distributed with zero mean. Also, backtransformation is needed when a reservoir simulation is run. If the distribution is not well defined in these transformations there will be a loss of information. In practice, transformations are done by table lookup and interpolation. The assumption of perfect knowledge about the distribution means that this table has a sufficiently high resolution to avoid loss of information. This may be a serious drawback for the gradual deformation method. However, it is important to recognize that the gradual deformation method is based on the conservation of statistical properties. If such statistical knowledge does not exist the gradual deformation method may not be the best choice from the start.

Table 5.1: Specifications of the wells. The wells are placed in a quarter-nine-spot pattern. Only water is being injected.

Well	Type	i	j	Constraint
I1	Inj	1	50	Rate ($5m^3/day$)
I2	Inj	1	50	Rate ($5m^3/day$)
I3	Inj	50	1	Rate ($5m^3/day$)
P1	Prod	50	50	BHP ($10bar$)

A reference field is generated by sequential Gaussian simulation and the corresponding production data is simulated and used as reference data. The reference permeability field is depicted in Figure 5.4(a). The reference data serves as fictive measurements of the production and is the data to be matched. The reference and all the realizations in the gradual deformations are generated with sequential Gaussian simulation with identical statistical parameters and are conditioned to 4 samples of permeability - one at each well location.

10 realizations are used in the matching, i.e. 10 outer cycles are performed. In each cycle a total of 10 function evaluations are allowed to find the optimal deformation parameter, yielding an overall of 100 objective function evaluations. Evaluation of the objective function involves one reservoir simulation to find the production data misfit.

Figure 5.2 shows the matched production data and the reference data. A good match is observed with only small deviations from the reference. In Figure 5.3 the evolution of the objective function is depicted. The figure shows a rapid decrease of the objective function in the first ten evaluations. Hereafter, the objective function decreases at a somewhat lower rate. The slowdown of the

convergence rate is a serious drawback for the gradual deformation method and happens because the parameterization used by the method is too restrictive on the search space used for the minimization.

Figure 5.4(b) shows the matched permeability field. The matched field reassembles the heterogeneity seen in the reference. A more rigorous test of the consistency of the match can be done using a qq-plot as depicted in Figure 5.4(c). The figure shows the quantiles of the reference versus the quantiles of the matched permeability field. If the the qq-plot is linear it is an indication that the two fields are distributed similarly (Petruccelli, Nandram & M.Chen 1999). The qq-plot does not appear particularly linear. However, an approximate linear trend can be observed. This indicates that the matched permeability approximately represents the reference heterogeneity. Another reason for the deviation from the straight line may be because of the relatively small number of grid blocks. Such a small number of data may not be sufficient to reproduce the statistical properties found in the reference which gives rise to ergodic fluctuations.

5.2.1 Application to a larger problem

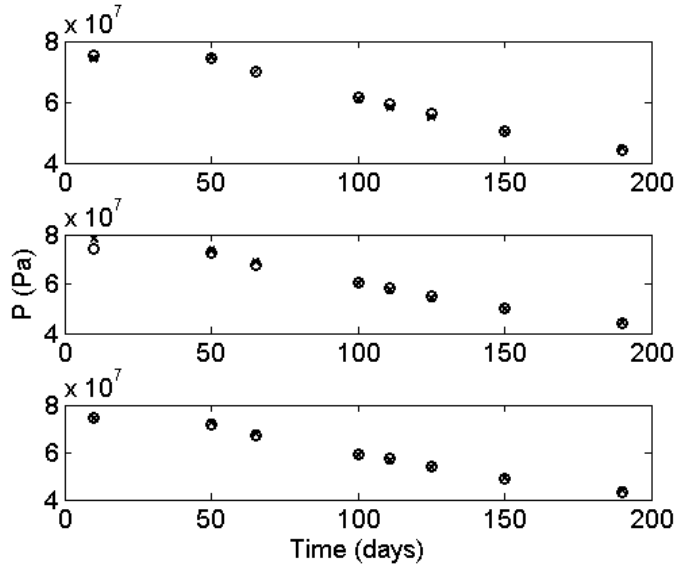
The previous example was based on a very coarse reservoir model with only 100 grid blocks and should be regarded as instructive rather than practical. To evaluate the use of the gradual deformation method on problems of a more realistic size a larger system will be introduced in the following discussion.

The reservoir we will deal with here is still defined in a two dimensional space. However, the reservoir is now discretized into a 100×100 grid, i.e. we have 10000 grid blocks to deal with. Again, the work is limited to an incompressible two phase system with water and oil present.

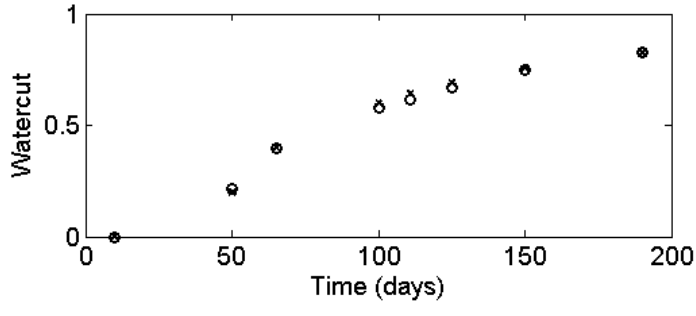
Figure 5.5 shows a permeability field with a 100×100 grid generated by sequential simulation. This field will serve as the reference in the following discussion. The reference shows a high correlation in the North-Eastward direction as a result of a non-isotropic variogram. Two water injectors are placed in the field along with 5 producers. The location of the wells is given in Table 5.2. The porosity is constant all over the reservoir and is set to 0.25.

Table 5.2: Specifications of the wells.

Well	Type	i	j	Constraint
I1	Inj	50	50	Rate ($5m^3/day$)
I2	Inj	50	85	Rate ($5m^3/day$)
P1	Prod	5	5	BHP ($10bar$)
P2	Prod	25	75	BHP ($10bar$)
P3	Prod	90	90	BHP ($10bar$)
P4	Prod	45	5	BHP ($10bar$)
P5	Prod	90	20	BHP ($10bar$)



(a) Injector pressure.



(b) Watercut

Figure 5.2: Match of injector pressures and watercut in a quarter-nine spot setup with the gradual deformation method. Ten realizations have been used in the matching process. Perfect knowledge about the permeability distribution has been assumed in the transformations involved with sequential Gaussian simulation. × : Reference, ○ : Match.

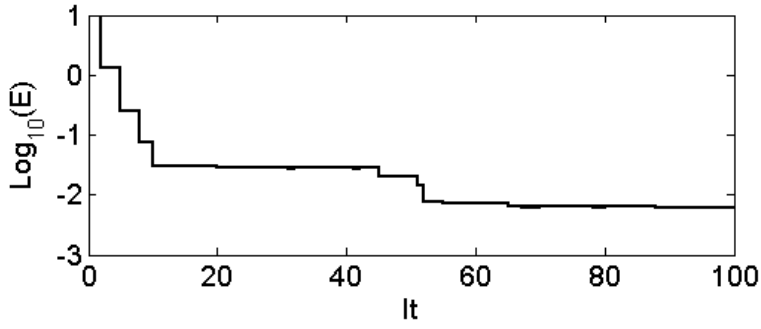
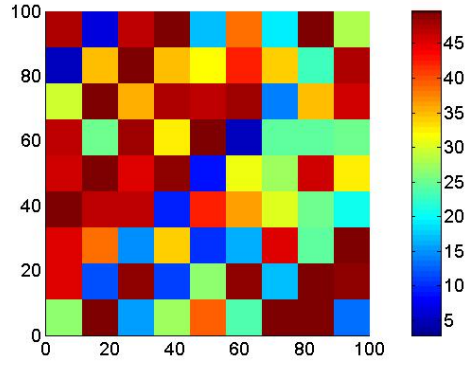


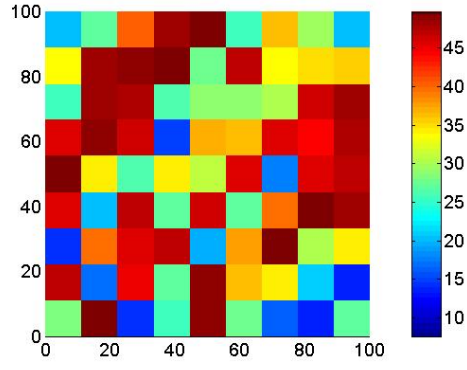
Figure 5.3: Evolution of the objective function during matching of the quarter-nine spot case with the gradual deformation method. Ten realizations are used and ten inner iterations are allowed in the Dekker-Brent routine.

Gradual deformation was applied to obtain a match of watercut and oil rates. Figure 5.7 shows the resulting permeability field and Figure 5.8 shows the watercuts at the production wells. The evolution of the objective function is shown in Figure 5.6. In more than 60 function evaluations the objective function is reduced by approximately 35 pct.. This rather poor performance must be attributed to the complexity of the problem. The well configuration makes history matching difficult because the individual wells influence each other. In other words, an improvement of the match at one well may lead to worsening of the match at other wells. This problem may be avoided by application of a suitable zonation strategy (Hoffman & Caers 2003). The qq-plot in Figure 5.9 indicates that the gradual deformation has conserved most of the statistics from the reference. However, a small deterioration can be seen. This may be due to accumulation of ergodic fluctuations from the deformations. The problem of accumulation of ergodic fluctuations in connection with gradual deformation is discussed in Hu (2002) where the more stable formulation involving three realizations for gradual deformation is introduced. This formulation has better numerical properties. However, the example just shown indicates that some drift in the distribution is still possible. Histograms of the reference permeability and the matched permeability are shown in Figure 5.10 . The histograms show that the distributions are approximately similar. Consequently, the deviation indicated by the qq-plot is not severe.

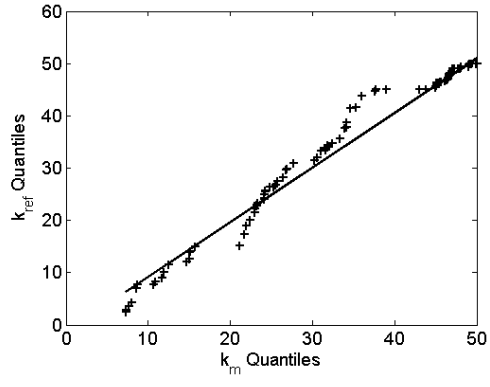
The use of the gradual deformation method in this dissertation is limited to reservoir models which can be represented by continuous variables. In Hu & Jenni (2005) the gradual deformation method is extended to object-based reservoir models. The gradual deformation method is therefore not limited to Gaussian fields but can be applied to more complex geologies if the geology can be represented by a Gaussian distribution at some level.



(a) Reference permeability.



(b) Best match from gradual deformation.



(c) qq-plot.

Figure 5.4: Reference permeability field and the field corresponding to the smallest production data mismatch in gradual deformation. The qq-plot is linear if the reference permeability and matched permeability are distributed similarly.

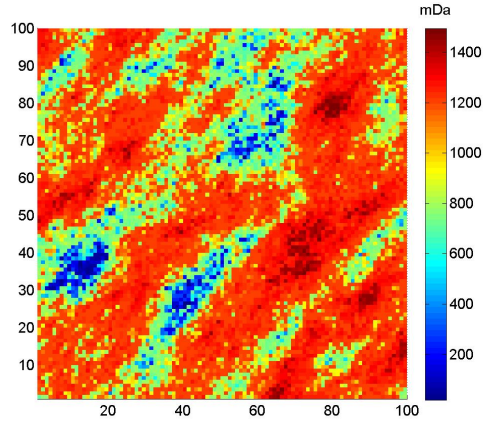


Figure 5.5: 100×100 reference field generated by sequential Gaussian simulation. A trend in the North-East direction is induced by a non-isotropic variogram.

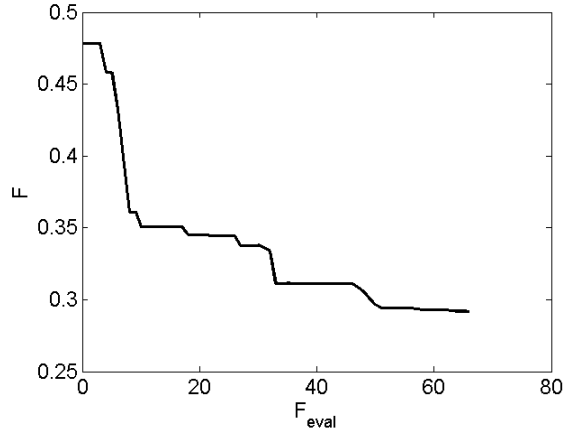


Figure 5.6: Evolution of the objective function during gradual deformation.

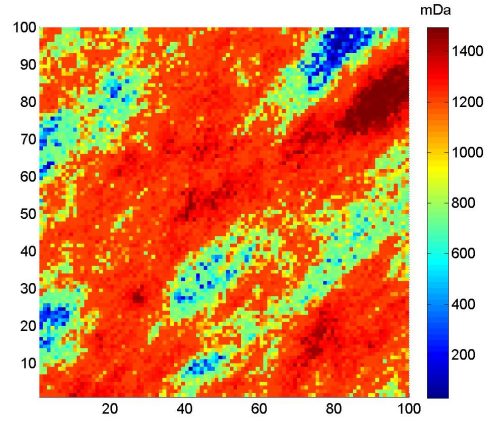


Figure 5.7: Resulting permeability after gradual deformation.

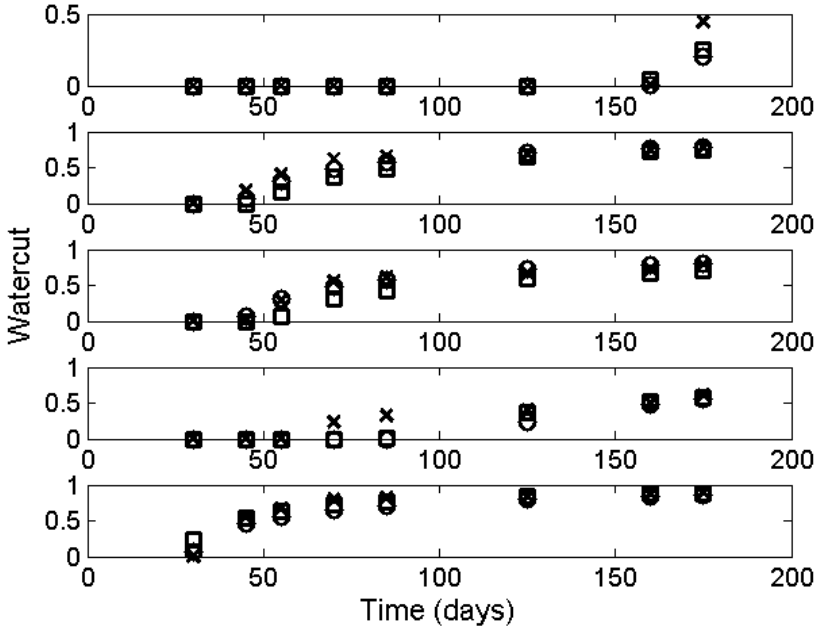


Figure 5.8: Watercut for the matched 180 days. × : Reference, ○ : Match, □ : Initial.

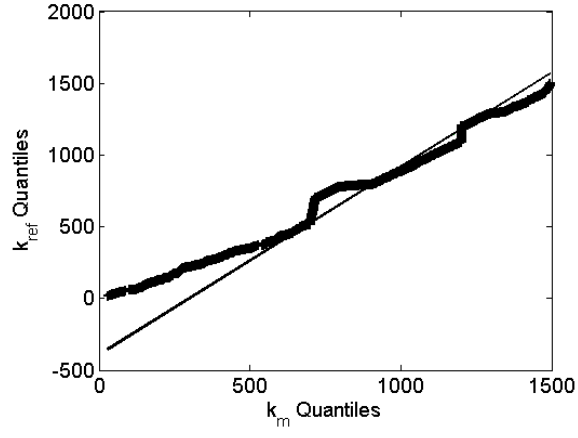
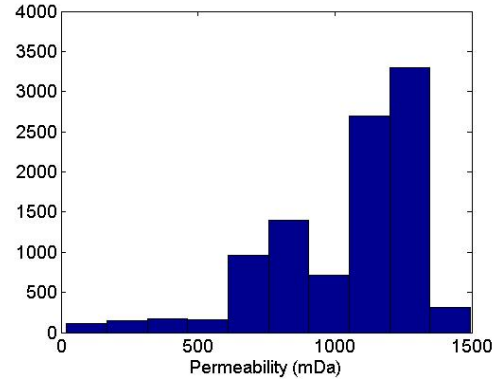


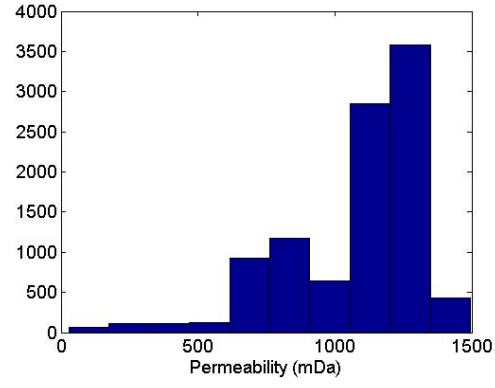
Figure 5.9: qq-plot of the quantiles of the reference permeability field versus the quantiles in the matched permeability field.

5.2.2 Regional gradual deformation

The gradual deformation method may be a too restrictive parameterization if the reservoir geology is complex or if the well configuration makes it difficult to match production data from one or more wells. One way of adding more flexibility to gradual deformation is to use more realizations to form the gradual deformation. This will involve more deformation parameters which means that history matching is not a one-parameter optimization problem anymore. The complexity of the optimization problem increases with the number of deformation parameters. Therefore, the number of deformation parameters should be defined as the best compromise between complexity and flexibility of the resulting optimization problem. Hoffman & Caers (2003) suggest a strategy which adds flexibility to history matching problems without increasing the complexity of the history matching problem. In their work the probability perturbation method makes up the framework for history matching. However, their proposed strategy can be applied to other methods because of it's simplicity and in this dissertation it has been applied in conjunction with gradual deformation. The main idea is to divide the reservoir into separate subregions and perturb these subregions independently from each other. The area between a producer and injector is an obvious choice to perturb in order to match the production data from the concerned wells. Hoffman & Caers (2003) suggests the use of streamline-derived regions to define subregions of the reservoir. Each zone is then adjusted independently from the others in order to match production data from wells placed in the zone. The use of a zonation technique together with gradual deformation has some complications. Since each zone is modified independently there will be a risk that the spatial correlation of the adjusted



(a) Reference permeability.



(b) Best match from gradual deformation.

Figure 5.10: Histograms of the reference permeability field and the permeability associated with the best production data match.

property is destroyed at the borders of the zones. Ravalec-Dupin, Nætinger, Hu & Blanc (2001) introduced a methodology where the gradual deformation method is extended in a way such that correlation across zones is conserved. This is achieved by performing the gradual deformation on uncorrelated normal deviates and then imposing the spatial correlation in a subsequent step. In the following examples the problem of conserving spatial dependence across zones is neglected. The validity of the neglection depends on the specific geology and the characteristic correlation length of the system. In the following example where the permeability is simulated using sequential Gaussian simulation it is a reasonably valid assumption to neglect spatial correlation across the boundaries. In more complex cases, e.g. channelized reservoirs, the spatial continuity across boundaries is more important. The case considered here is only aimed at demonstrating some of the advantages of a zonation approach and is by no means exhaustive.

The following example is based on a symmetry element of a quarter-nine spot well configuration. The quadratic reservoir is discretized into a 100×100 grid. Reservoir simulation is done using the streamline simulator 3DSL (Batycky 1997). The use of a streamline simulator makes it possible to define streamline-derived regions and the computational load is reduced compared to conventional finite difference based simulators. In this case the definition of the regions is straightforward since the injectors are only connected to one producer. Each injector/producer pair defines a drainage area in which the injected water will flow. The main idea behind the zonation approach is that the drainage area associated with one injector has a larger impact on the data from the particular injector than other parts of the reservoir. Thus, the geological properties of the drainage area will be adjusted in order to match production data at the associated injector. Since only one producer is present production data from the producer can not be attributed to only one drainage area since the producer is connected to several injectors. Injector bottomhole pressures and producer watercut are matched in the following example.

The following three-dimensional objective function is defined:

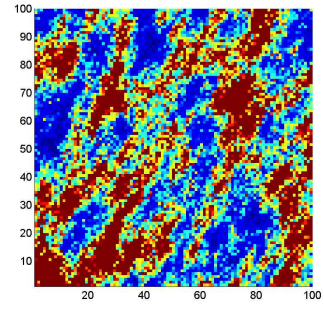
$$E_k = \sum_{i=1}^{N_T} \left(\frac{p_i^{Sim} - p_i^{Obs}}{p_i^{max}} \right)_k^2 + (WCUT_i^{Sim} - WCUT_i^{Obs})_k^2, \quad k = \{1, 2, 3\}. \tag{5.10}$$

Watercut is part of all three elements of the objective function since the producer data can not be attributed to a single drainage area. Figure 5.11(a)-(d) illustrate how the regions are defined. The streamline simulator traces streamlines going from injectors to producer. From the streamlines the borders defining separate zones can be derived. In the illustrated case separatrix lines are used to define the borders. Alternatively, the streamlines may be used directly to define the regions. In the latter case there will be parts of the reservoir which

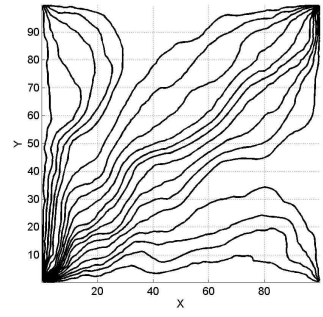
are not associated with a injector/producer pair. In more complex cases with high heterogeneity and a more random well placement the use of separatrix lines to define regions may be difficult to realize. Finally, the regions are defined according to the borders. Each region is now associated with one of the elements in the objective function given in equation (5.10). One important detail must be stressed: Even though the objective function is multidimensional and the reservoir is divided into subregions only *one* reservoir simulation is carried out per iteration during history matching. If an element of the objective function is decreased during history matching the gradual deformation of the associated region is accepted even if the deformation leads to an increase in one of the other elements. The main assumption of the strategy is that the other wells are influenced very little when a region is modified.

Figure 5.12 shows the evolution of the three objective functions. The three objective functions all decrease at different iterations. If global deformation had been used the gradual deformation might not have been able to improve the match because improvement of one matched property may lead to worse matches at other wells.

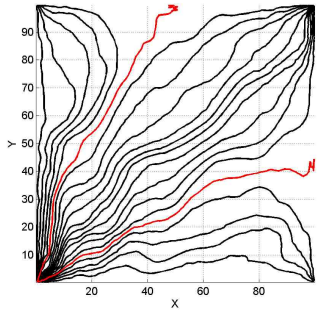
The example demonstrates some of the advantages of zonation. However, for more complex scenarios it may by difficult to define regions because of high interdependency of the wells. Another issue arises from the non-stationary nature of multiphase flow in porous media. The placement of the streamlines may change considerably over time. If this is the case it is not obvious how to choose representative regions. The use of zonation techniques is not discussed any further in this dissertation. Because of time constraints and more importantly restraints on the software available it has not been possible to investigate the use of zonation any further.



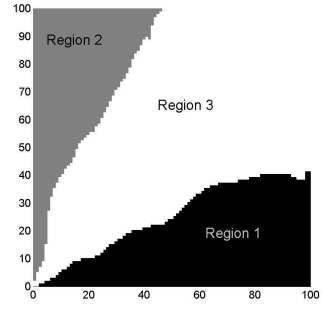
(a) Field.



(b) Trace streamlines.



(c) Define borders.



(d) Define regions.

Figure 5.11: Possible workflow for the computation of separate regions. The streamlines are traced by a streamline simulator. In this work the 3DSL simulator was used (Batycky 1997).

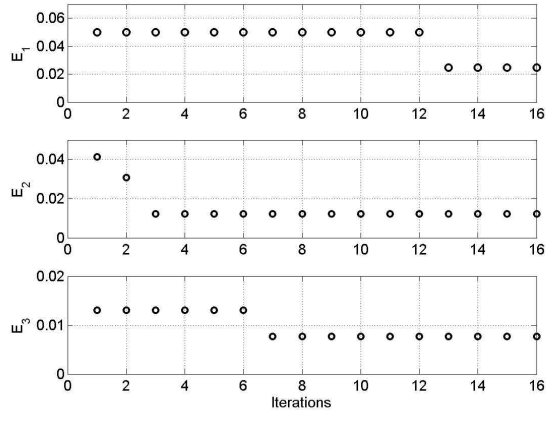


Figure 5.12: Evolution of the objective functions for the three regions during matching. Note that the objective functions decrease at different iterations.

Chapter 6

Probability perturbation

One of the main reasons for developing geostatistically based history matching methods is that such methods make it easier to constrain the reservoir model to certain geological properties. Often, geologic knowledge comes in the form of conceptual models of the geology of the reservoir. The conceptual model may be built from outcrops or previously developed fields and many other sources. Since the conceptual model is build from analogous fields it is qualitative and serves as a qualified guess on the type of geology found in the reservoir. The conceptual model is not necessarily conditioned to other data such as samples from the wells but is qualitative. Combination of qualitative information with quantitative knowledge from samples taken at well locations and seismic surveys is not an easy job because of the diversity of the information sources. The emergence of the training image based sequential simulators has made it possible to generate realizations of complex geological scenarios conditioned to hard data. The realizations can also be conditioned to soft data from e.g. seismic data. In applications related to history matching the realizations have to honor production data as well. Caers (2003) introduces a method known as the probability perturbation method (PPM), which exploits the algorithmic structure of sequential simulation. PPM is a highly flexible data integration method which can be applied to history matching problems where the geology is complex. In this chapter the background of the method is presented and some sample applications are shown. The method is extended to include qualitative gradient information in Chapter 8.

6.1 Introduction to the probability perturbation method

The probability perturbation method was introduced by Caers (2003) as a method for data integration into constrained geologic models. PPM exploits the fact that sequential simulations make use of local constrained probability functions when unknown properties are simulated. When a grid note with undefined properties is visited during sequential simulation a local probability function is constructed from the data associated with the neighboring grid notes. The local probability function will be denoted $P(A|B)$ where A denotes an event, e.g. "channel facies present" and B denotes the data in the neighboring grid notes. The probability is constrained to the data B. The way the probability is computed differs from algorithm to algorithm. In Gaussian simulation $P(A|B)$ is found by solving a local kriging system and in Snesim it is found by traversing a search tree derived from a training image. Since the simulated value at a grid note is found from a local probability function, secondary information can be introduced into the simulation if it can be represented in terms of probability.

Secondary information could be seismic information used to condition a realization of porosity or permeability. $P(A|C)$ will denote the probability function given the secondary information, C.

A piece of secondary information has to be combined with the conditional probability function $P(A|B)$ when a property is drawn at a grid block. Combining two conditional probabilities can be done by the following expression, known as the τ-model (Journel 2002):

$$\frac{x}{a} = \left(\frac{b}{a}\right)^{\tau_1}\left(\frac{c}{a}\right)^{\tau_2}, \tag{6.1a}$$

where

$$x = \frac{1-P(A|B,C)}{P(A|B,C)} \tag{6.1b}$$

$$b = \frac{1-P(A|B)}{P(A|B)} \tag{6.1c}$$

$$c = \frac{1-P(A|C)}{P(A|C)} \tag{6.1d}$$

$$a = \frac{1-P(A)}{P(A)}. \tag{6.1e}$$

$P(A)$ is the global probability of the event A occurring, also denoted the *marginal distribution* of A. In PPM the typical values for the exponents are $\tau_1 = \tau_2 = 1$ (Caers 2007).

For $\tau_1 = \tau_2 = 1$, rearrangement of equation (6.1a) gives

$$P(A|B,C) = \frac{1}{1+x} = \frac{a}{a+bc} \tag{6.2}$$

The idea behind PPM is to integrate the production data mismatch in history matching into the generation of new realizations. This is done by treating the mismatch as a piece of secondary information. PPM uses the following parameterization to translate the production data mismatch into secondary information:

$$P(A|C) = (1 - r_c)\boldsymbol{i} + r_c P(A), \tag{6.3}$$

where r_c is a perturbation parameter and $\boldsymbol{i}$ is a realization of binary indicator variables for a certain physical property at the grid blocks. C now represents the production data. In equation (6.3) it is assumed that only two indicators can represent the geology. This corresponds to the situation where the reservoir consists of two facies. The binary indicator variable i consists of 0's and 1's which indicate the type of facies at the grid blocks. $P(A)$ is the marginal distribution of the event A. The marginal distribution is the unconditional probability of A and is usually the global proportion of A over the entire domain. When r_c in equation (6.3) is equal to 0 the conditional probability will be equal to the binary indicator, i.e. $P(A|C) = \boldsymbol{i}$. In this case a sequential simulation conditioned to $P(A|C)$ will result in a realization identical to $\boldsymbol{i}$. If r_c takes the value 1 the conditional probability is reduced to the unconditional probability and a sequential simulation will result in a new binary indicator variable equiprobable to $\boldsymbol{i}$. If r_c is between 0 and 1 the result will be a binary indicator which is a mix between $\boldsymbol{i}$ and another equiprobable realization. Thus, the parameter r_c can be used to perturb the binary indicator variable. Analogously to gradual deformation the geology of a reservoir can be perturbed systematically by adjusting r_c.

It is important to note that the conditional probability does not depend on the production data mismatch, C. Therefore, it is intuitively better to consider the production mismatch to be dependent on the resulting realization of binary indicators. The production data therefore depends on the choice of r_c. Consequently, r_c is a free parameter which can be adjusted in order to improve the history match. Since $P(A|C)$ as given in equation (6.3) is independent of the production data mismatch, the probability perturbation method is a stochastic search algorithm. The production data mismatch is included into the simulation through $P(A|C)$ as a soft probability constraint. This means that the result from the simulation will honor statistical properties such as the variogram in the case of Gaussian simulation and multiple-point statistics when training image based simulators are applied.

PPM and gradual deformation both make use of parameterizations which make it possible to constrain history matches to certain parameters. Where gradual deformation is limited to geologies which can be represented by a Gaussian distribution[1], PPM is more flexible since no assumptions regarding the

[1] In Hu & Jenni (2005) the gradual deformation method is extended to object-based reservoir models.

statistics of the indicator variable have been made. This makes PPM a good choice for matching of complex geological scenarios. The similarity with gradual deformation is also seen in the algorithm used for history matching which also makes use of an inner loop to search for the optimal value for r_c and an outer loop over the seeds used for sequential simulation.

The PPM algorithm can be summarized as:

1. Generate an initial realization of binary indicators conditioned to available hard data and a training image
2. Change the seed for the sequential simulator
3. Define $P(A|C) = (1 - r_c)\boldsymbol{i}^{(k)} + r_c P(A)$
4. In an inner loop find the optimal r_c and the corresponding realization $\boldsymbol{i}_{r_c}^{(k)}$
5. If the objective function decreased, update $\boldsymbol{i}^{(k+1)} = \boldsymbol{i}_{r_c}^{(k)}$. Else set $\boldsymbol{i}^{(k+1)} = \boldsymbol{i}^{(k)}$
6. $k = k + 1$
7. If not matched and $k < k_{max}$, go to 2

Caers (2003) uses the Dekker-Brent method to search for the optimal r_c. In the implementation of PPM used in this dissertation the Dekker-Brent method has also been applied. This is due to the discontinuous behavior of the objective function when the reservoir geology is binary and has curvilinear features. In such cases a derivative-free method can be expected to be more feasible than traditional derivative based algorithms. The perturbation parameter, r_c, is not required to be the same over the entire domain of interest. In Hoffman & Caers (2003) regional probability perturbation is suggested. In their work different zonation techniques are combined with PPM in order to increase the flexibility of the method. The use of streamline-derived regions to divide the reservoir into separate regions is shown to increase the capabilities of PPM. If the properties of some regions of the reservoir are more certain than others it may be desired to keep those properties relatively unchanged through history matching. This can be achieved by keeping the perturbation parameter close to 0 in the regions with high certainty.

The fact that PPM does not require that the adjusted parameter can be represented by a certain distribution makes the method highly versatile. Consequently, PPM can be used as the engine in a history matching setup which can integrate information from very diverse sources. Such information may be a qualitative description of the geology supplied by a geologist together with hard data samples at well locations and a seismic survey. The conceptual model from the geologists can be used as a training image for sequential simulations conditioned to the hard data and soft data if it is available. Thus, a priory

knowledge on the reservoir geology is integrated into the reservoir modelling directly through conditioning and training images for the sequential simulator. Finally, the production data can be integrated through an optimization procedure with PPM as the main engine for the production data integration. The history matched reservoir will reassemble the geological scenario depicted in the training image and honor sampled values as well as secondary information which has been input as soft constraints. Figure 6.1 illustrates the structure of a PPM based history matching procedure.

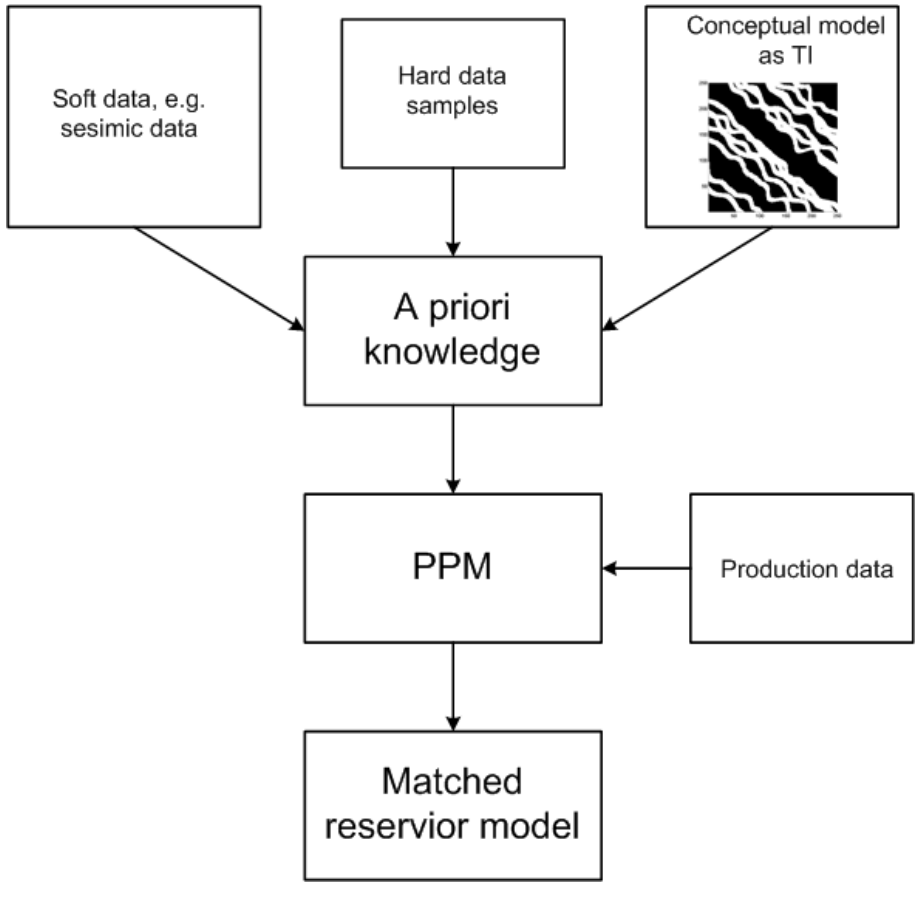


Figure 6.1: Principal structure of an algorithm utilizing probability perturbation to constrain history matches to various a priori data including a conceptual model of the reservoir geology.

6.1.1 The role of the training image

The local conditional distribution functions used in sequential simulation can be derived from a training image depicting a desired geology. In the Snesim algorithm a search-tree is built from the training image as a preprocessing step. Whenever a local distribution function is calculated the algorithm runs through the search-tree and looks for patterns similar to the constellation of points in the vicinity of the simulated point. Thus, the nature of the training image controls the overall properties of the simulation result. The training image may be a pixelized hand-drawing from a geologist or a pixelized areal photo of a process analogous to the process that created the reservoir, e.g. a fluvial delta. A training image may also be derived from seismic explorations or from combinations of the mentioned sources. One of the main advantages of the use of training image based simulators in conjunction with PPM is that history matching can be performed on complex geological scenarios without deteriorating the geological consistency of the reservoir model. In the paper of Caers (2003) the PPM is introduced and in order to keep the applications simple and clear it is assumed that the training image is known. Of course, this is an approximation since the geologic scenario is also an uncertain factor in real field applications. Handling of uncertain geological scenarios is discussed in Suzuki & Caers (2006) where a distance measure is used to search for training images which are likely to improve the history match. Coupled to PPM this methodology is able to handle uncertainties in the geological scenario in a history matching setup.

6.2 Application of PPM to history matching examples

This section illustrates the use of PPM on a number of simple but instructive cases. The first case involves the channelized reservoir depicted in Figure 6.2. The reservoir has been generated by an unconditional simulation in Snesim and will serve as the reference field to be matched by the PPM procedure. The reservoir consists of two facies, one is high-permeable (e.g. sandstone) and forms the channels from which the oil is produced. A low-permeable (e.g. mud) facies separates the high-permeable zones. The field is operated with two water injectors and three producers placed in a pattern, which will result in a sweep of injected water along the channels. The locations of the wells are given in Table 6.1. The high-permeable facies has a permeability of $1000mDa$ and the low-permeable facies has a permeability of $1mDa$. The porosity is assumed to be 0.25 over the entire domain. The field is discretized with a $40 \times 40 \times 1$ equidistant Cartesian grid. When the field is history matched it is assumed that the only unknown property is the location of the facies. The permeabilities of the facies are assumed to be known in advance and the Snesim realizations are conditioned to facies data at the well locations. History matching will be carried

out for injector bottomhole pressure, oil rates, and watercuts at the producers. Measurements of the production data are picked out at 11 times in order to have representative data for the dynamic behavior of the system. Ten outer iterations are allowed in PPM. The inner iterations are allowed to perform a maximum of five function evaluations. Note that one function evaluation corresponds to running a full reservoir fluid flow simulation.

The objective function used for the quantification of the production mismatch is:

$$E = \sum_{j=1}^{N_{inj}} \sum_{i=1}^{N_T} \left(\frac{p_{i,j}^{Sim} - p_{i,j}^{Obs}}{p_{max}^{Obs}} \right)^2 + \sum_{j=1}^{N_{prod}} \sum_{i=1}^{N_T} (WCUT_{i,j}^{Sim} - WCUT_{i,j}^{Obs})^2, \quad (6.4)$$

where N_T is the number of times where production data is measured. N_{inj} and N_{prod} are the number of injections wells and production wells. p_{max}^{Obs} is the maximum pressure in the reference data.

6.3 shows the evolution of the objective function versus the number of function evaluations. In about 45 function evaluations PPM has improved the objective function with more than two orders of magnitude. In Figures 6.4, 6.5, and 6.6 the matched production data is shown. The figures show the production data for the realization used to initialize the history matching procedure and the production data for the matched field and the reference. Note that the pressure in the second injector for the initial realization has been omitted because it is much higher than the matched pressure and the reference pressure. Inclusion of the initial pressure would change the scale of the pressure axis. The figures show that the PPM method is able to improve the production data match significantly within a reasonable number of function evaluations. Figures 6.7(a) and 6.7(b) depict the initial realization used to initialize the history matching and the final matched field, respectively. The main difference between the two fields is that the connection between injector 2 and producer 3 has been restored in the matched field. The same overall heterogeneity is seen in the reference and the matched field because the two fields are derived from the same training image.

Table 6.1: Specifications of the wells.

Well	Type	i	j	Constraint
I1	Inj	25	20	Rate ($5m^3/day$)
I2	Inj	5	2	Rate ($5m^3/day$)
P1	Prod	38	4	BHP ($10bar$)
P2	Prod	13	38	BHP ($10bar$)
P3	Prod	2	34	BHP ($10bar$)

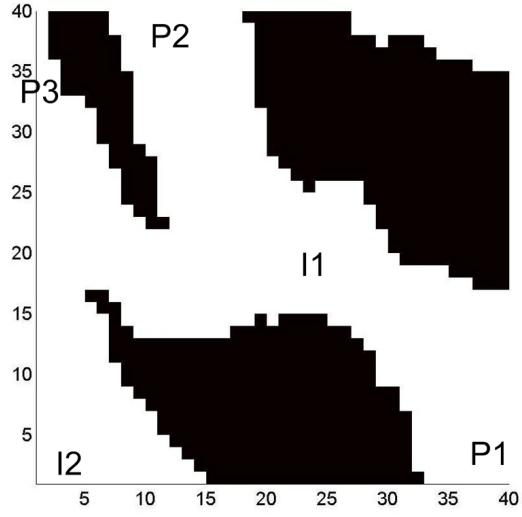


Figure 6.2: Reference permeability field. White indicates facies 1 with permeability $1000mDa$ and black indicates the low permeable (1mDa) facies 2.

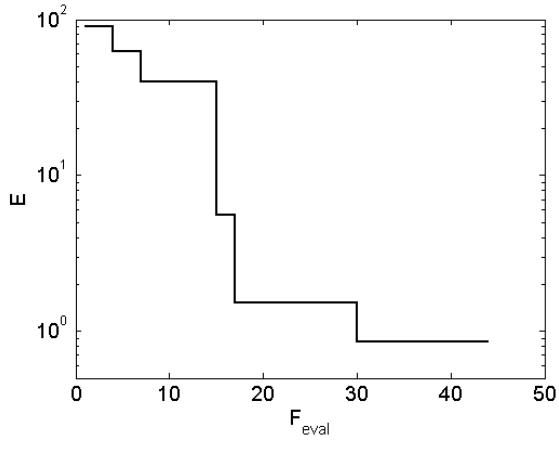


Figure 6.3: Evolution of the objective function during history matching. 10 outer iterations are allowed.

6.2.1 Multiple history matches

PPM is a stochastic history matching method and may converge to different results depending on the initial field used to initiate history matching with. The seeds used in the subsequent simulations of the facies also affect the final result. Consequently, PPM can be used to obtain multiple history matches. An ensem-

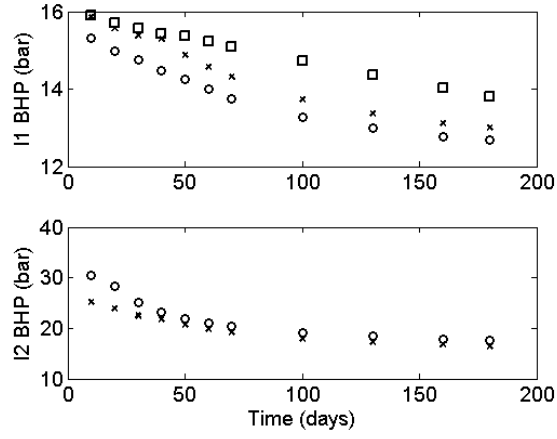


Figure 6.4: Bottomhole pressures for the two injectors. □ : Initial, ○ : Match, × : Reference. The initial pressures in injector 2 are very high and are not included in the lower graph.

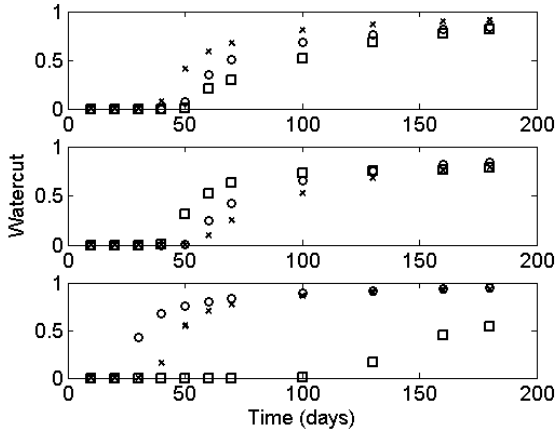


Figure 6.5: Watercuts for the three producers. □ : Initial, ○ : Match, × : Reference.

ble of history matched fields can be used to quantify the uncertainties in the reservoir model and the forecasts of the production data. The field considered before is now matched several times. This results in an ensemble of permeability fields which are conditioned to the same hard data and training image and which have been adjusted to honor the reference production data. The first case with multiple history matches will take its starting point in the reference field in Figure 6.2 with the same well configuration as before. Again, the production data is measured at eleven representative times. The measurements include the

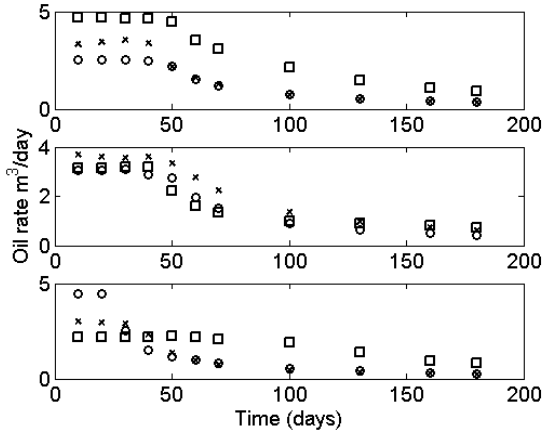


Figure 6.6: Oil rates for the three producers. □ : Initial, ○ : Match, × : Reference.

late-time behavior of the reservoir. Therefore, the results are not applicable for validation of the predictive capabilities of the reservoir models. Later, a more realistic case, where only the early behavior of the reservoir is matched, will be presented. 49 matches have been found by PPM. The evolution of the 49 objective functions is depicted in Figure 6.8. From the figure it is evident that not all 49 cases lead to good matches. If a threshold of 2.0 is chosen as a cutoff for the objective function value only 21 cases will be retained. The 21 best realizations are summarized in Figure 6.9 which shows the so-called E-type of the field. The E-type contains the mean of the facies type indicators at each grid block, i.e.

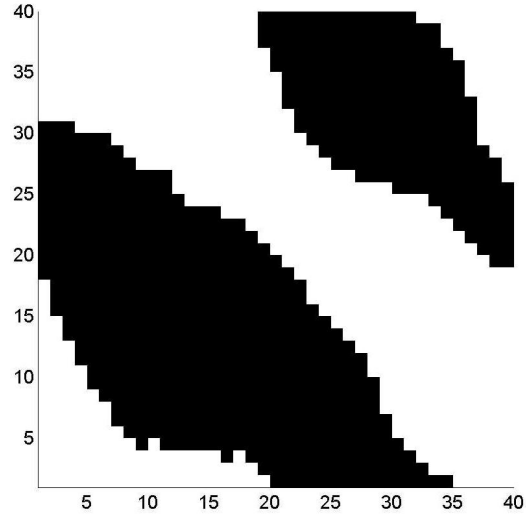
$$E = \frac{1}{N_{rel}} \sum_{j=1}^{N_{rel}} \boldsymbol{i}_j, \tag{6.5}$$

where N_{rel} denotes the number of retained realizations.

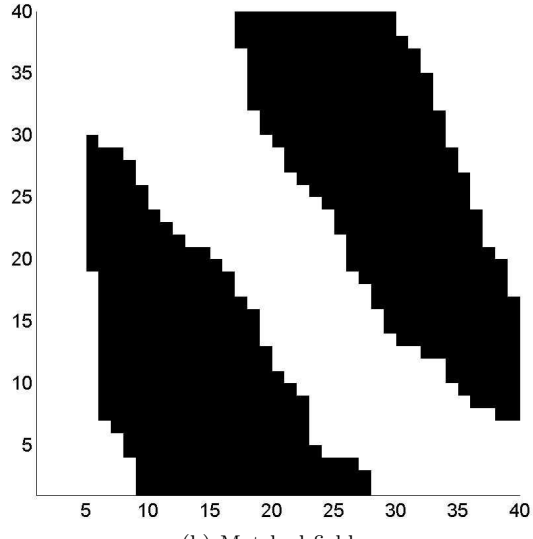
Light areas in the E-type indicate areas where the certainty of finding the high permeable sand facies is high (indicator 0), darker areas indicate where it is more certain to find mud facies (indicator 1). The E-type is in good agreement with the reference which is mainly because the generated realizations are derived from the same training image as the reference. In addition, the certainty is increased by the inclusion of hard data from the well locations.

Figures 6.10, 6.11, and 6.12 show the mean of the production data from the 21 matched fields. 95 pct. confidence intervals are indicated by the error bars and the reference data is marked with circles [2]. Most of matched production data lies within or close to the 95 pct. confidence intervals. However the matched

[2]Confidence intervals are made under the assumption that the data is normally distributed.



(a) Initial field.



(b) Matched field.

Figure 6.7: Initial permeability field and matched permeability field. The initial field is used to initialize the PPM algorithm.

pressure in injector 2 lies systematically over the reference pressure. This implies that the area close to this injector should be modified if a better match should

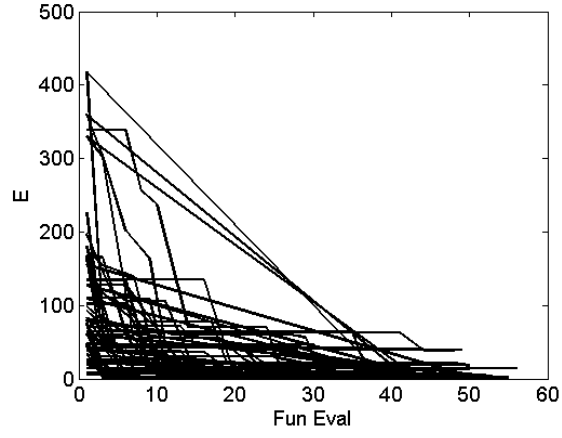


Figure 6.8: Evolution of the 49 objective functions during history matching.

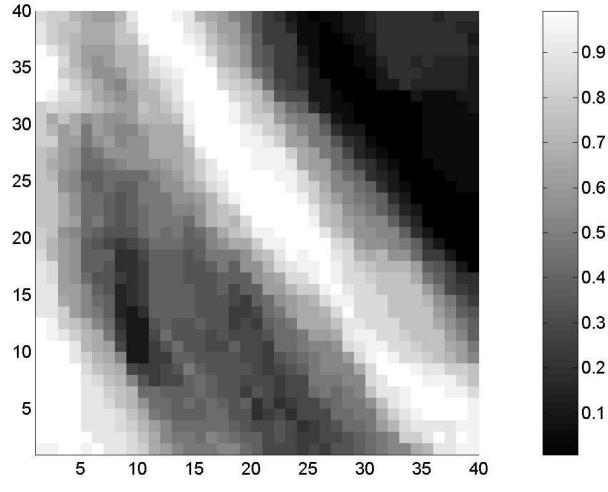
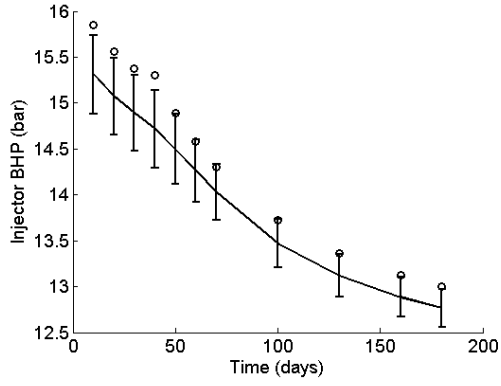
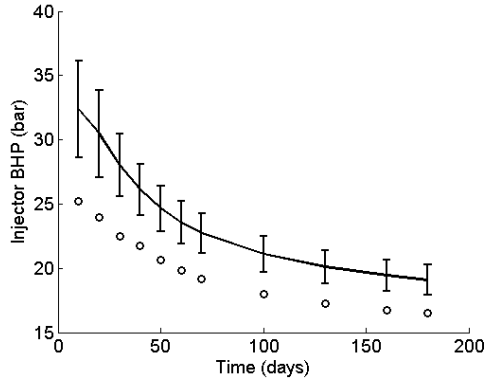


Figure 6.9: E-type of 21 fields which resulted in an objective function value less than 2. White areas indicate areas with a high certainty of finding high permeable sand facies and dark ares indicate areas with a high certainty of finding low permeable mud.

be obtained. The size of the error bars depends on the spread of the matched production data. One way of reducing the spread is to increase the number of matched fields and to set a lower cut-off for when to accept the match.



(a) Injector 1.



(b) Injector 2.

Figure 6.10: The solid line shows the mean of the 21 best matches of injector pressures. Circles indicate the reference pressures. The error bars indicate a 95 pct. confidence interval for the matched production data.

Prediction with multiple matched fields

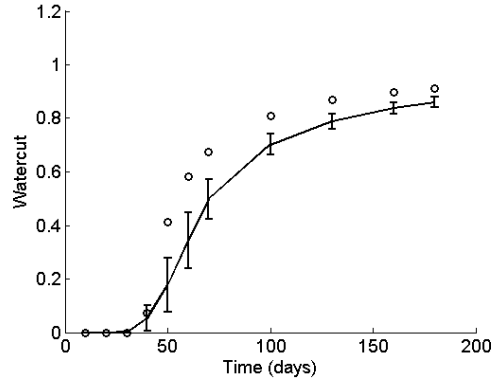
One of the most important reasons for history matching the production of an oil field is to use the calibrated reservoir model to predict future production. As illustrated above, the PPM method can be used to generate multiple history matched fields. From the spread of the predicted production data the uncertainty of the predictions can be assessed. The field shown in Figure 6.2 will be used as the reference field again but only the first six measurements of production data will be matched. This corresponds to matching the first 60 days of production. The matched fields will then be used to predict the production until

180 days of production. A total of 52 history matches have been run whereof 13 result in a final objective function value less than 2. Figures 6.13, 6.14, and 6.15 show the mean of the matches for the 13 best matches. The solid line indicates the matched data and predictions are indicated with a dashed line. The circles indicate the reference measurements of the production data and the error bars show a 95 pct. confidence interval for the predictions and the matches[3]. Nearly all of the measured production data lie within or close to the 95 pct. confidence intervals (except for pressure in injector 2 again). However, the intervals are wide witch indicate that the predictions are associated with a high uncertainty. This is a valuable piece of information which may prevent wrong decisions for future operations because of over-confidence in the reservoir model. Figure 6.16 shows the E-type field made from the 13 matches. The E-type resembles the E-type from the case where the full production history was matched (Figure 6.9). This is mainly because the fields are outcomes of the same process; they are conditioned to the same hard data and training image. The spread of the predicted watercuts seem to decrease in the late time predictions. This is counter intuitive since the spread of the predictions should increase as the prediction horizon increases. However, the physics play an important role in the setup considered here. The watercut will eventually converge to 1 irrespective of the quality of the history match. This means that the late time behavior is less uncertain than the early behavior. This is an important detail to keep in mind when evaluating the quality of the matched reservoir model and its predictions of future production.

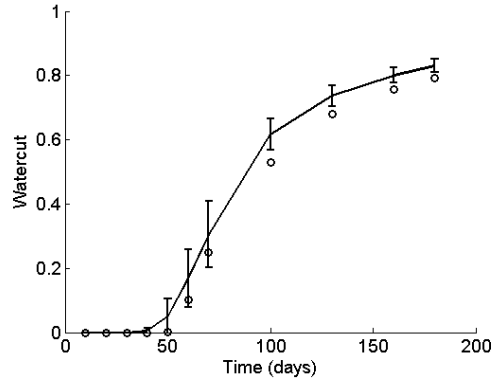
6.3 Summary of the probability perturbation method

The PPM method is a geostatistically based parameterization of the history matching problem. The method takes advantage of the algorithmic structure of sequential simulation where local distribution functions are used to simulate unknown properties. In a probability based framework the PPM method can be used to history match reservoir models with complex geological features. The method can be used to acquire multiple history matched models which can be used to asses the certainty of predicted production. One drawback of the method is that it requires many reservoir simulations before an acceptable match is obtained. This is due to the stochastic nature of PPM where gradient information is not applied for the optimization.

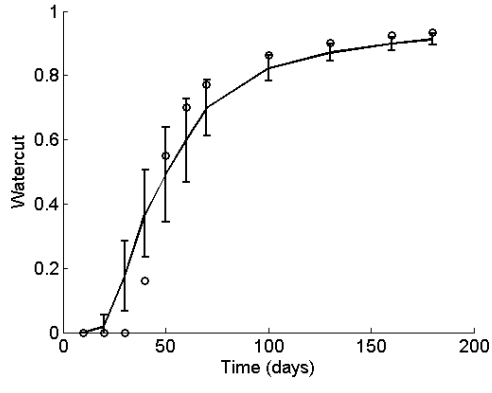
[3]The data is assumed to be normally distributed around the reference data. It could be argued that 13 predictions are insufficient to quantify uncertainty.



(a) Producer 1.

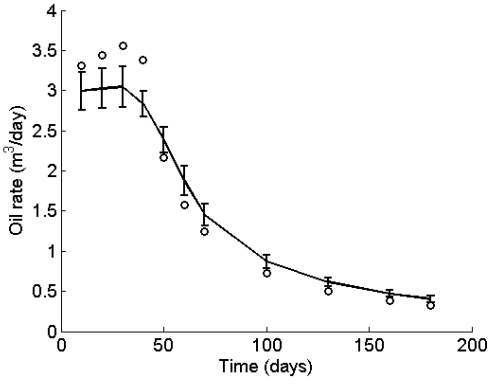


(b) Producer 2.

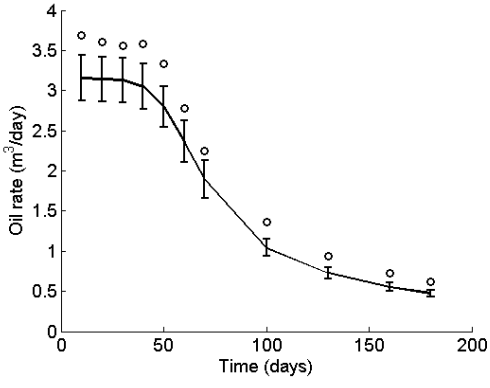


(c) Producer 3.

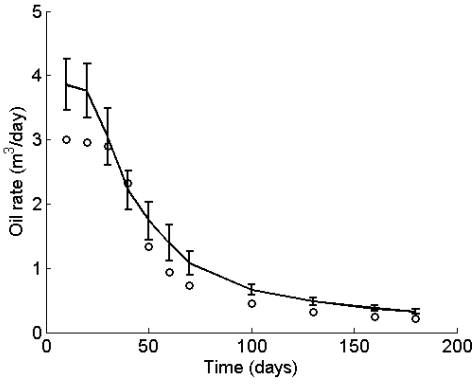
Figure 6.11: The solid line shows the mean of the 21 best matches of watercut. Circles indicate the reference. The errorbars indicate a 95 pct. confidence interval for the matched production data.



(a) Producer 1.

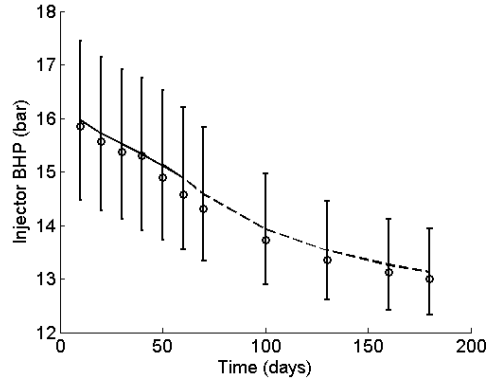


(b) Producer 2.

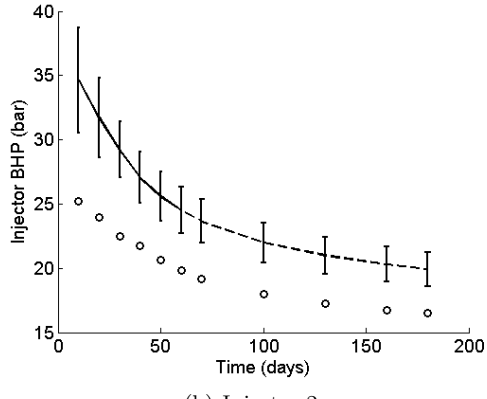


(c) Producer 3.

Figure 6.12: The solid line shows the mean of the 21 best matches of oil rate. Circles indicate the reference. The errorbars indicate a 95 pct. confidence interval for the matched production data.

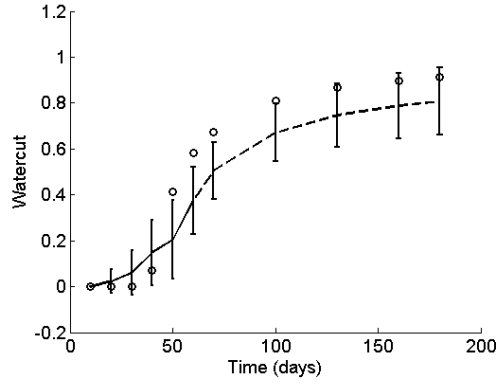


(a) Injector 1.

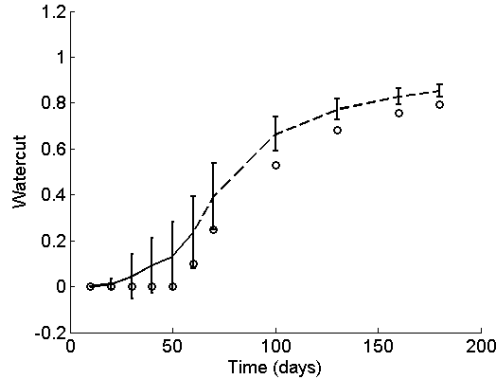


(b) Injector 2.

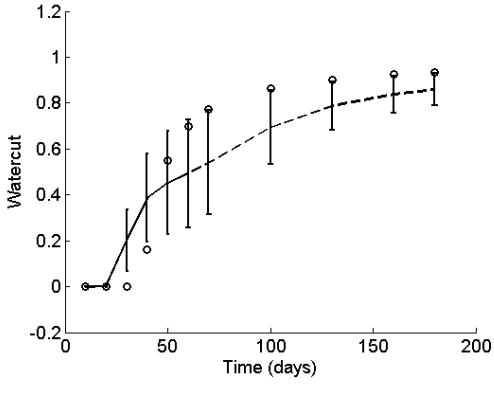
Figure 6.13: The solid line shows the mean of the 13 best matches of injector pressures. The predicted pressure is indicated with the dashed line. Circles indicate the reference pressures. The error bars indicate a 95 pct. confidence interval for the matched production data.



(a) Producer 1.

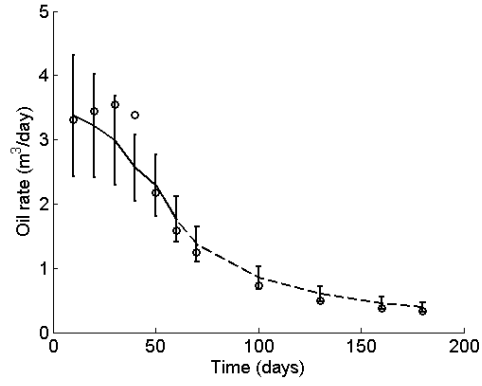


(b) Producer 2.

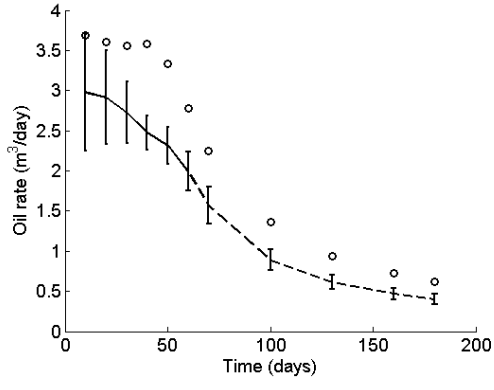


(c) Producer 3.

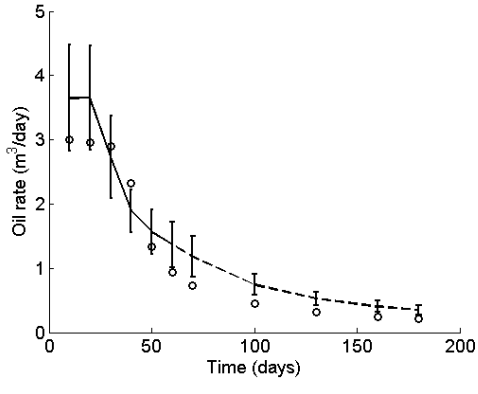
Figure 6.14: The solid line shows the mean of the 13 best matches of watercut. The predicted watercut is indicated with the dashed line. Circles indicate the reference. The error bars indicate a 95% confidence interval for the matched production data.



(a) Producer 1.



(b) Producer 2.



(c) Producer 3.

Figure 6.15: The solid line shows the mean of the 13 best matches of oil rate. The predicted oil rate is indicated with the dashed line. Circles indicate the reference. The error bars indicate a 95 pct. confidence interval for the matched production data.

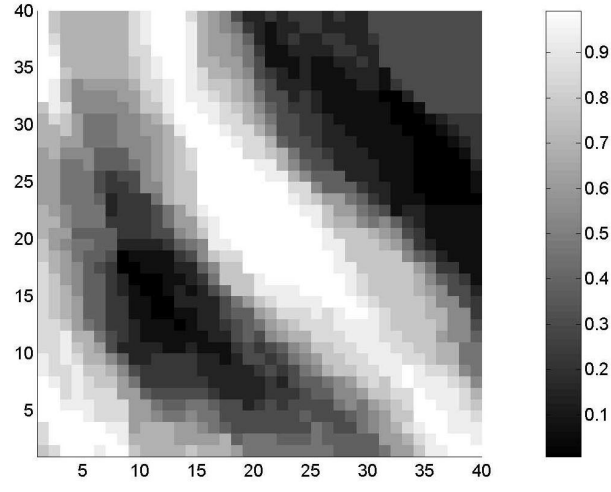


Figure 6.16: E-type of 13 fields which resulted in an objective function value less than 2. White areas indicate areas with a high certainty of finding high permeable sand facies and dark ares indicate areas with a high certainty of finding low permeable mud.

Chapter 7

Deterministic history matching

The history matching procedures described until now have been stochastic. In a geostatistical framework history matching has been parameterized and simplified. Methods like PPM and gradual deformation make it easy to constrain the match to statistical properties and the underlying optimization problem is much simpler than the multi-dimensional problem encountered in the "raw" history matching problem. The main drawback for the stochastic methods is their slow convergence toward a match. This is due to the complexity of history matching of oil production. Another reason for the poor convergence are the severe parameterizations which reduce the search space for the adjusted parameters significantly. An alternative to the stochastic methods are the deterministic methods which are based on well-known optimization algorithms. Where the stochastic methods search for an improved reservoir model randomly, the deterministic methods make use of derivative information when updating the reservoir model. Thus, a deterministic history matching procedure requires an efficient calculation of the sensitivities of the matched production data with respect to the adjusted grid block properties. The fact that sensitivities with respect to grid block properties are required has made the use of deterministic methods in history matching computationally demanding. However, the emergence of the adjoint method for the calculation of sensitivities and improved capabilities of hardware have enabled the use of derivative-based methods on large field cases. This chapter introduces a deterministic method and sample applications are demonstrated. An integral part of the method is regularization of the history

matching problem which is used to reduce the intrinsic ill-posedness of history matching and to ensure some geological consistency of the result. Regularization will also be discussed in the present chapter.

7.1 Introduction to deterministic methods for history matching

The term *deterministic methods* covers all the classical derivative-based optimization methods such as steepest descent methods, conjugate gradient methods, and the well-known methods for nonlinear least squares problems: Gauss-Newton and Levenberg-Marquardt. In this dissertation the choice of method has fallen on the Levenberg-Marquardt method. The Levenberg-Marquardt method is specially designed for nonlinear least squares problems which is exactly what history matching problems often are formulated as. A brief introduction to the theory behind the method will be presented along with a few important implementational details. Wu et al. (1998) and Li et al. (2003) both discuss automatic history matching by the use of nonlinear least-squares methods. These methods are pointed out as being more efficient than simpler techniques such as the steepest descent method.

The Levenberg-Marquardt method is explained in standard textbooks on optimization methods, e.g. Nocedal & Wright (1999). However, in order to make the following discussion more clear for readers without any experience with methods for non-linear least-squares problems, a short description of the Levenberg-Marquardt method is provided in Appendix C.

The sum-of-squares objective function is defined as

$$E = \frac{1}{2} \sum_{i=1}^{N_{obs}} w_i (d_i^{obs} - d_i^{sim})^2, \tag{7.1}$$

where d denotes production data. The observed and simulated production data is collected in the vectors $\boldsymbol{d}^{obs}$ and $\boldsymbol{d}^{sim}$, respectively. Minimization of the sum-of-squares objective function is analogous to minimizing the 2-norm of $\boldsymbol{d}^{obs} - \boldsymbol{d}^{sim}$, i.e. minimization can be formulated as:

$$\boldsymbol{m}^* = \underset{\boldsymbol{m}}{\text{Argmin}} \left[\|\boldsymbol{d}^{obs} - \boldsymbol{d}^{sim}\|^2 \right], \tag{7.2}$$

where $\boldsymbol{m}$ represents the adjustable parameters, e.g. grid block permeabilities. Negative values of permeability and porosity are prohibited since these parameters are non-negative per definition. This may cause problems when the history matching problem is stated as an unconstrained optimization problem and solved by a deterministic approach. To avoid negative values of the adjusted

parameters it is convenient to apply a logarithmic transformation and optimize with respect to a transformed variable:

$$\log \mathbf{k}^* = \underset{\log \mathbf{k}}{\text{Argmin}} \left[E(\mathbf{k}) \right] \tag{7.3}$$

The gradient with respect to the transformed variable is calculated as:

$$\frac{\partial E}{\partial \log \mathbf{k}} = \mathbf{k} \frac{\partial E}{\partial \mathbf{k}}. \tag{7.4}$$

This transformation will be used in the deterministic approach to avoid negative permeabilities during the optimization.

If N_{tot} measurements of production data are available the residual can be defined as:

$$\boldsymbol{r} = \boldsymbol{d}^{obs} - \boldsymbol{d}^{sim} = \begin{bmatrix} d_1^{obs} - d_1^{sim} \\ d_2^{obs} - d_2^{sim} \\ \vdots \\ d_{N_{tot}}^{obs} - d_{N_{tot}}^{sim}. \end{bmatrix} \tag{7.5}$$

The Jacobian of the residual vector is defined as:

$$\boldsymbol{J}_{i,j}^r = \frac{\partial r_i}{\partial m_j}, \tag{7.6}$$

where r_i is the ith element of the residual and m_j is the jth independent variable. The gradient of the objective function is then given as

$$\boldsymbol{g} = \frac{\partial E}{\partial \boldsymbol{m}} = -(J^r)^T \boldsymbol{r}. \tag{7.7}$$

In the following discussion the term *sensitivity* will denote the derivative of a measurable quantity at the wells with respect to the grid block properties, i.e.

$$\boldsymbol{s}_i = \frac{\partial d_i^{sim}}{\partial \boldsymbol{m}}. \tag{7.8}$$

The adjoint procedure presented in Section 4.3 provides the sensitivities of the production data with respect to grid block porosities and permeabilities.

History matching is an inverse problem and is usually ill-posed. The ill-posedness arises because the measured quantities do not carry sufficient information to characterize the reservoir description fully. This means that different reservoir descriptions may lead to equally good matches. Consequently, the objective function can be expected to exhibit an irregular behavior and to be composed of many local minima. Another problem arises from the fact that the number of measurable production data is smaller than the number of adjustable parameters in the history match. This makes the history matching under-determined which hinders the use of many efficient optimization methods.

One way of dealing with ill-posed and under-determined problems is to apply a regularization strategy. In the following section the concept of regularization is introduced.

7.1.1 Regularization in history matching

The sum-of-squares objective function takes the general form:

$$E = \frac{1}{2}\sum_{i=1}^{N_{obs}} w_i (d_i^{obs} - d_i^{sim})^2. \tag{7.9}$$

As mentioned above the objective function can be expected to have many local minima and exhibit a highly irregular behavior in most history matching problems. One way of making the objective function behave more smoothly is to apply regularization. A widely used method to regularize ill-posed problems is *Tikhonov regularization*. A general Tikhonov regularization augments the objective function with a regularization term which penalizes deviations from a prior set of independent variables (Tikhonov & Arsenin 1977), which in general history matching terms translates to:

$$E = \frac{1}{2}\sum_{i=1}^{N_{obs}} w_i (d_i^{obs} - d_i^{sim})^2 + \frac{\sigma_r^2}{2}\sum_{i=1}^{N_{ind}} (\boldsymbol{m} - \boldsymbol{m}_0)^2, \tag{7.10}$$

where σ_r is the regularization parameter which controls the impact of regularization.

The regularization term stabilizes the ill-posed optimization problem because it has a smoothly varying and continuous behavior when $\boldsymbol{m}$ is changed. Obviously, the regularization term increases when the changes to the independent variable become larger.

Wu et al. (1998) and Li et al. (2003) make use of a similar augmented objective function:

$$E(\boldsymbol{m}) = \frac{1}{2}((\boldsymbol{m}-\boldsymbol{m}_{prior})^T \boldsymbol{C}_M^{-1}(\boldsymbol{m}-\boldsymbol{m}_{prior}) + (\boldsymbol{g}(\boldsymbol{m})-\boldsymbol{d}_{obs})^T \boldsymbol{C}_D^{-1}(\boldsymbol{g}(\boldsymbol{m}))-\boldsymbol{d}_{obs}), \tag{7.11}$$

where $\boldsymbol{m}$ denotes the set of unknown reservoir parameters (vertical and horizontal permeability and well skin factors). Matrices $\boldsymbol{C}_M$ and $\boldsymbol{C}_D$ denote covariance matrices of the model parameters and observed data, respectively. The variance of production data is calculated in different ways depending on the type of data. Measurement errors for pressures and gas oil ratios are calculated as independent Gaussian distributions and errors on water oil ratios are scaled with the magnitude of the measurement. The operator $\boldsymbol{g}$ denotes the process of running a reservoir simulation. One of the advantages of penalizing deviations from a prior set of independent variables is that prior information on the geology can by

input through the regularization term. This ensures some geological consistency of the match.

An augmented residual is defined as

$$\boldsymbol{r} = \boldsymbol{d}^{obs} - \boldsymbol{d}^{sim} = \begin{bmatrix} d_1^{obs} - d_1^{sim} \\ d_2^{obs} - d_2^{sim} \\ \vdots \\ d_{N_{tot}}^{obs} - d_{N_{tot}}^{sim} \\ \sigma_r(m_1 - m_1^0) \\ \vdots \\ \sigma_r(m_{ind} - m_{ind}^0) \end{bmatrix}, \tag{7.12}$$

where the deviations from the prior set of variables are added to the vector of production mismatches. In the following applications of the gradient-based method the norm of the augmented residual is minimized. It can be difficult to define a suitable value for the regularization parameter from a priori knowledge. Therefore, some trial-and-error investigations where impacts of regularization are tested may be useful before the actual history matching is initiated. As the production mismatch decreases during optimization the regularization term will become more and more dominant in the augmented objective function. Consequently, the convergence of the optimization method will deteriorate as the match becomes better. A dynamic update of the regularization parameter may by used to reduce the impact of the regularization term as the production match is improved. Haber, Ascher & Oldenburg (2000) suggest a simple continuation strategy where the regularization parameter is gradually reduced when the regularization term becomes too large compared to the other components of the objective function. Such a strategy has been implemented in the methods used in this dissertation.

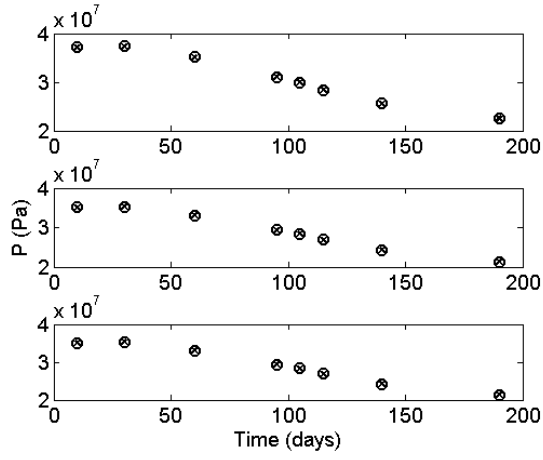
The regularization term keeps a handle on what the optimizer does to the reservoir model since extreme values are penalized harder. Therefore, the optimizer will generally seek to change many independent variables a little bit instead of changing a few variables by a large amount. This results in a smoothing effect and the resulting sets of independent variables can be expected to behave smoothly. This may often be contrary to the geological expectation where sharp contrasts in the reservoir properties are often seen. In a binary facies model one would not expect smoothly varying properties and the use of a gradient-based history matching method may not be appropriate. As shown in Section 7.5 gradient based methods are not always suitable for a matching of particular reservoirs; especially categorical facies models pose a challenge for gradient based methods.

7.2 Practical application of the gradient-based history matching techniques

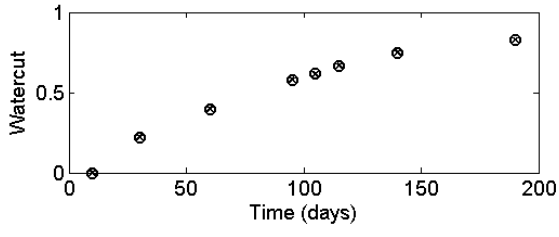
7.2.1 First simple example

The first case involves a simple quarter-nine-spot well configuration. The reference permeability field is shown in Figure 7.5(a) and is generated by sequential Gaussian simulation. Production takes place in the upper right corner and water injection is taking place in the three other corners of the reservoir. Porosity is assumed constant in the entire domain and is assumed to be known. Permeability is known at the four well locations. A homogeneous field of $50mDa$ is chosen as prior. Thus, deviations from $50mDa$ are penalized by the regularization term. Figures 7.1(a) and (b) show the matched production data. The figure shows that the production data matches very well with the reference. However, if the matched permeability field in Figure 7.5(b) is compared with the true permeability in Figure 7.5(a) it is evident that the match is far from the truth. This is an example of the non-uniqueness of the history matching problem. A field with significantly different features than the true field is still capable of producing production curves which are very close to the reference. This circumstance shows that the prior field used as initial guess and in the regularization is of high importance for the result. Figure 7.2 shows the evolution of the objective function. The objective function is split into two parts, one corresponding to the squared production mismatch and the second corresponding to the regularization term. After 3 iterations the production data term becomes smaller than the regularization term and hereafter no improvement of the production mismatch occurs.

To illustrate the importance of the prior permeability the same reference is now matched but with a different prior. The prior is now chosen to be a realization from the same process that generated the reference. The prior is now conditioned to the four samples and honors the second order statistics found in the reference. The matched production data is seen in Figure 7.3. Figure 7.4 shows the evolution of the objective function. Again, the match is obtained after 3-4 iterations and improvement stops when the regularization term exceeds the production mismatch term. The resulting permeability field is shown in Figure 7.5 (c). As expected, the resulting match exhibits a less smooth behavior than in the case where the prior field was homogeneous. All this indicates that the prior permeability field has a very large impact on how the resulting permeability field is distributed.



(a) Injector pressure.



(b) Watercut

Figure 7.1: Match of injector pressures and watercut in a quarter-nine spot setup with the Levenberg-Marquardt method using Tikhonov regularization. The prior field is homogeneous with permeability 50 mDa. × : Reference, ∘ : Match.

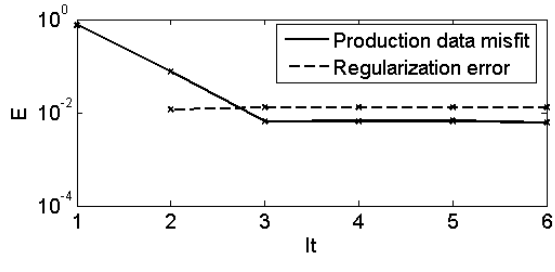
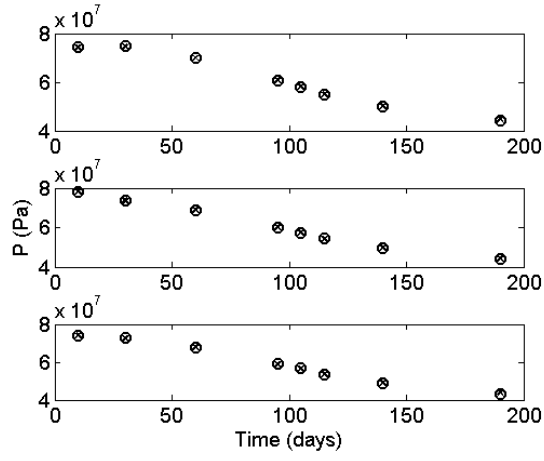
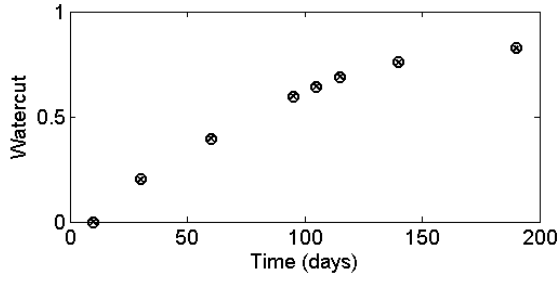


Figure 7.2: Evolution of the objective function during matching of the quarter-nine spot case with the Levenberg-Marquardt method with a homogeneous prior field. The objective function is split into a production data mismatch part and a regularization part.



(a) Injector pressure.



(b) Watercut

Figure 7.3: Match of injector pressures and watercut in a quarter-nine spot setup with the Levenberg-Marquardt method using Tikhonov regularization. The prior field is heterogeneous and comes from the same random process which was used to generate the reference field. × : Reference, ∘ : Match.

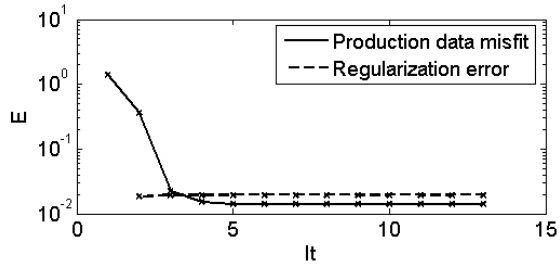


Figure 7.4: Evolution of the objective function during matching of the quarter-nine spot case with the Levenberg-Marquardt method with a heterogeneous prior field. The objective function is split into a production data mismatch part and a regularization part.

7.3 Combination of deterministic history matching and stochastic methods

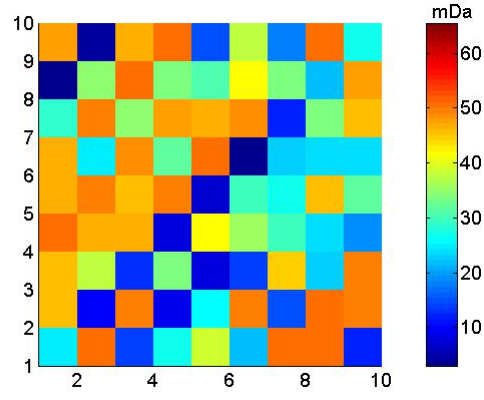
As we saw in the previous section the prior permeability field has a great influence on the result. Even with a prior with significantly different properties than the reference the Levenberg-Marquardt method is able to match the production data well. If there is a priori information about the statistics of the permeability available it may only be passed to the Levenberg-Marquardt routine by the prior field in the current setup. To make use of the advantages of gradual deformation as well as benefitting from the efficiency of the deterministic method the following sequential approach to history matching is suggested:

1. By gradual deformation find a realization of permeability that matches the production data sufficiently well.

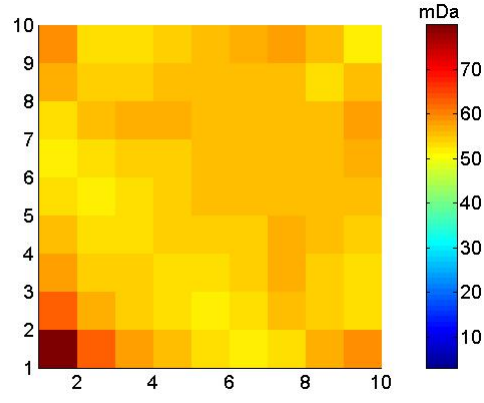
2. Adjust the gradual deformation result by the use of Levenberg-Marquardt with the gradual deformation result as the prior.

The sequential approach is also discussed in Johansen, Shapiro & Stenby (2006).

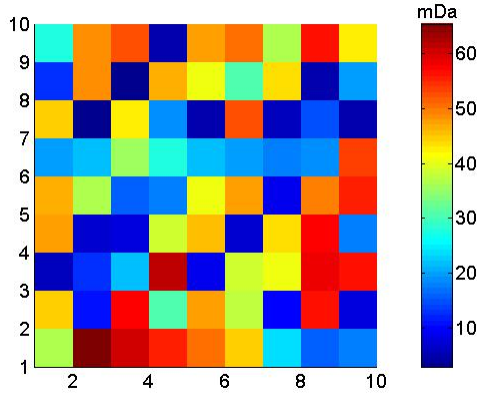
The proposed combination uses the deterministic method as a final adjustment of the permeability. If the match from the gradual deformation is sufficiently good this will not introduce large changes to the permeability. Thereby, the statistics of the prior is likely not to change significantly during the final tuning. The sequential approach is applied to the same history matching problem as before. Figure 7.6 shows the matched data. As in the other cases a good match is achieved. In figure 7.7 the objective function evolution during the final adjustment is depicted. A decrease of the objective function happens quickly and after 3-4 Levenberg-Marquardt iterations the match does not improve any further. The production mismatch and the regularization terms weigh approximately the same in the final iterations. The fact that the final value of the



(a) Reference.



(b) Prior homogeneous.

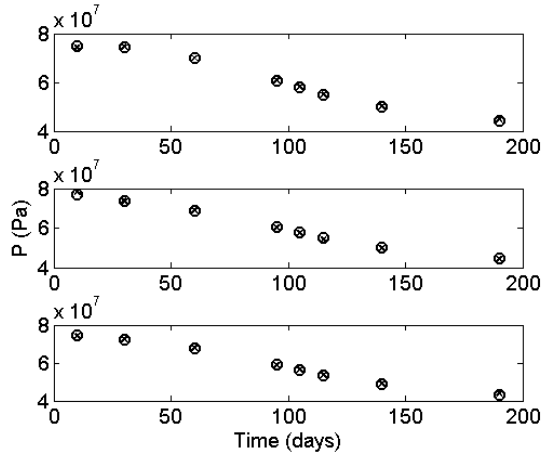


(c) Prior heterogeneous.

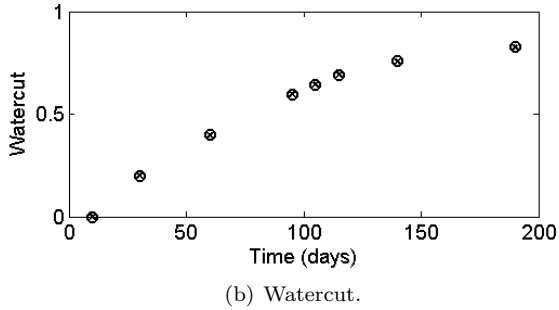
Figure 7.5: The reference field and the resulting fields after matching. The reference is generated by sequential Gaussian simulation and has a high correlation in the North-East direction. The reference has a mean permeability of $33.9mDa$, (b) has a mean of $56.0mDa$ and (c) has a mean of $32.2mDa$.

norm of the production mismatch is lower than for the other cases is encouraging but should not be interpreted as a direct measure of the appropriateness of the method. Since many (or infinitely many) realizations of permeability may result in perfect matches a low objective function alone is not guaranty for the reliability of the result.

Figure 7.8(a) shows the result from the gradual deformation which is used as the prior field in the Levenberg-Marquardt method. The final result after adjustment is shown in Figure 7.8(b).



(a) Injector pressure.



(b) Watercut.

Figure 7.6: Match of injector pressures and watercut in a quarter-nine spot setup with the Levenberg-Marquardt method using Tichonov regularization. The prior field is heterogeneous and comes from a gradual deformation process. × : Reference, ∘ : Match.

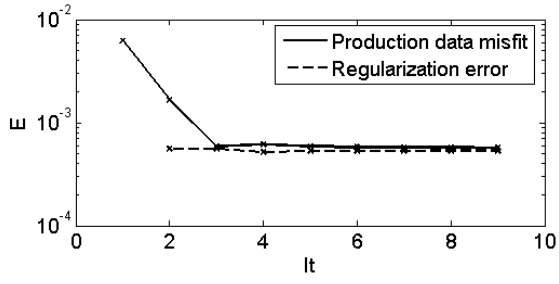
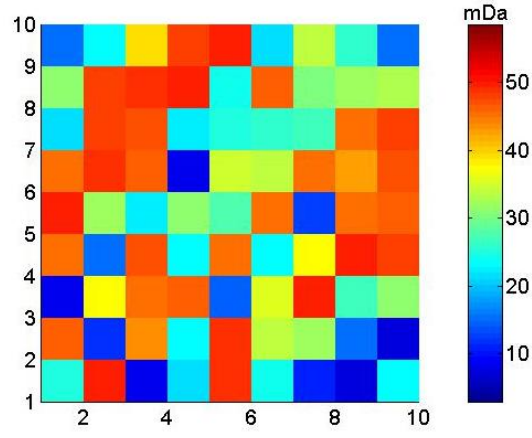
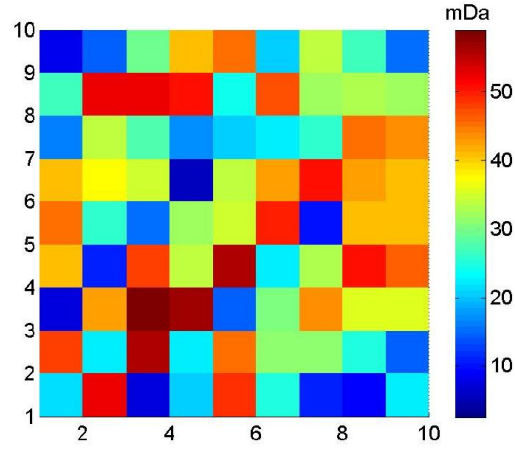


Figure 7.7: Evolution of the objective function during matching of the quarter-nine spot case with the Levenberg-Marquardt method with a heterogeneous prior field from the gradual deformation process. The objective function is split into a production data mismatch part and a regularization part.



(a) After gradual deformation.



(b) After adjustment.

Figure 7.8: Permeability fields for the case where the gradual deformation result is used as prior for the Levenberg-Marquardt routine. (a) is the outcome of the gradual deformation method. (b) is the final result. The mean permeability of (a) is 33.3mDa and the mean of (b) is 32.2mDa.

7.3.1 Application to a larger problem

To evaluate the use of the Levenberg-Marquardt method and gradual deformation on problems of a more realistic size a larger system will be introduced in the following discussion.

The reservoir we will deal with here is still defined in 2D. However, the

reservoir is now discretized into a 100×100 grid, i.e. we have 10000 grid blocks to deal with. The fluids (water and oil) are still incompressible.

Figure 7.9 shows a permeability field with a 100×100 grid generated by sequential simulation. This field will serve as the reference in the following discussion. The reference shows a high correlation in the North-Eastward direction as a result of a non-isotropic variogram. Two water injectors are placed in the field along with 5 producers. The location of the wells is given in Table 7.1. The porosity is constant all over the reservoir and is set to 0.25.

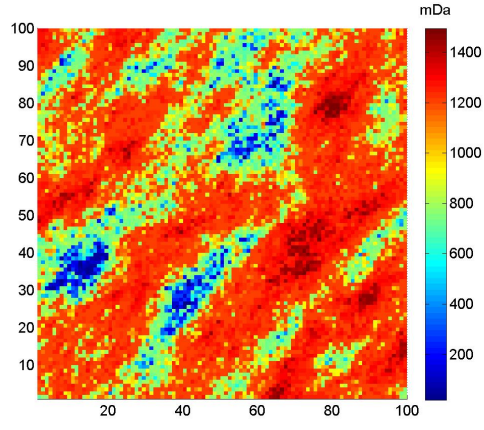


Figure 7.9: 100×100 reference field.

Table 7.1: Specifications of the wells.

Well	Type	i	j	Constraint
I1	Inj	50	50	Rate $(5m^3/day)$
I2	Inj	50	85	Rate $(5m^3/day)$
P1	Prod	5	5	BHP $(10bar)$
P2	Prod	25	75	BHP $(10bar)$
P3	Prod	90	90	BHP $(10bar)$
P4	Prod	45	5	BHP $(10bar)$
P5	Prod	90	20	BHP $(10bar)$

As shown in the previous discussion it may be an advantageous strategy to apply the gradual deformation method to find a prior permeability field for the deterministic approach. By starting out with gradual deformation we ensure that the prior is in better accordance with the measured production data and that the following adjustment by the deterministic method is smaller. Since

the final adjustment is likely to be small it is also likely that the second order statistics which were conserved in the gradual deformation will be more or less intact after the adjustment. Of course this approach relies on the assumption that gradual deformation is capable of performing a sufficiently large part of the matching on its own, i.e. that the objective function is brought to a sufficiently low level.

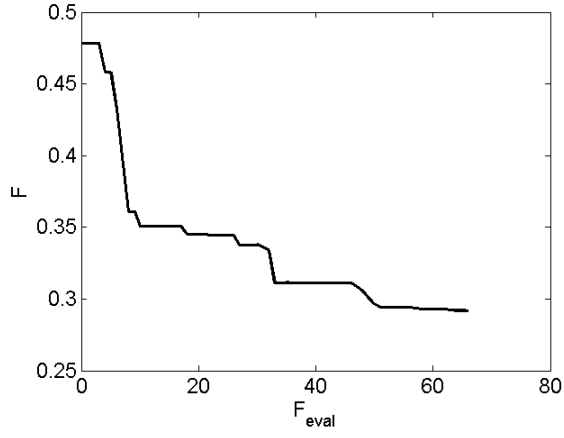


Figure 7.10: Evolution of the objective function during gradual deformation.

Gradual deformation was applied to obtain a match of watercut and oil rates. Figure 7.11 shows the resulting permeability field and Figure 7.12 shows the watercuts at the production wells. The evolution of the objective function is shown in Figure 7.10. In more than 60 function evaluations the objective function is reduced by approximately 35 pct.. This rather poor performance must be attributed to the complexity of the problem. The well configuration makes history matching difficult because the individual wells influence each other. In other words, an improvement of the match at one well may lead to its worsening at other wells. This problem may be avoided by application of a suitable zonation strategy (Hoffman & Caers 2003).

The result from gradual deformation is used as an initial model for the Levenberg-Marquardt method. To avoid an overcorrection of permeability near the wells the regularization parameter σ_r is scaled according to the kriging variances. The regularization parameter at grid block (i, j) is calculated by:

$$\sigma_{r,ij} = \frac{c}{((\sigma_{ij}^k)^2)^n + k}, \quad k \in]0; 1], \tag{7.13}$$

where $(\sigma_{ij}^k)^2$ is the kriging variance at the particular grid block. This ensures that the regularization becomes larger close to the wells because the kriging

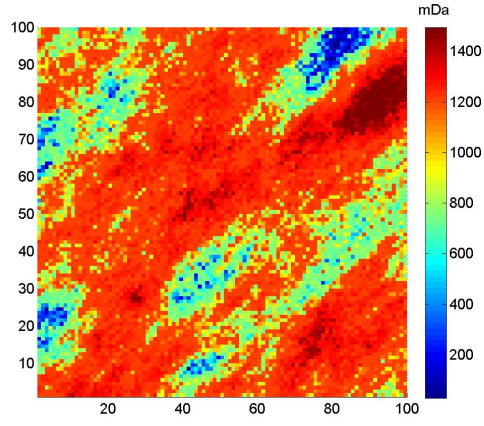


Figure 7.11: Resulting permeability after gradual deformation.

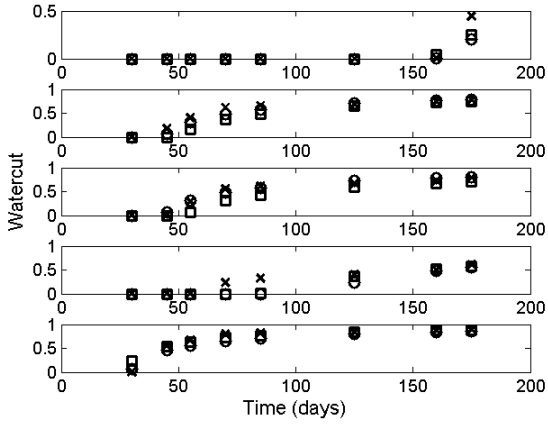


Figure 7.12: Watercut for the matched 180 days. × : Reference, ○ : Match, □ : Initial.

variance is zero at a sampled location and increases with the distance to the location. For small variances the regularization parameter has the following limiting behavior:

$$\sigma_{r,ij} \rightarrow \frac{c}{k}, \quad \text{for} \quad (\sigma_{ij}^{k})^2 \rightarrow 0. \tag{7.14}$$

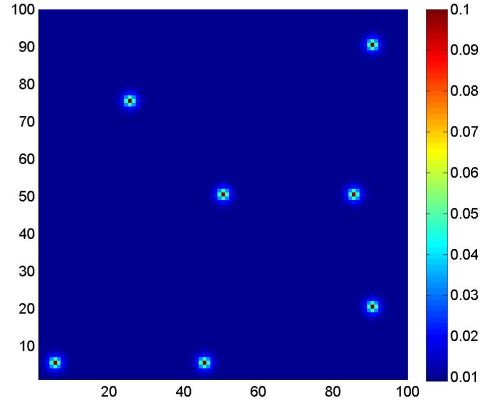


Figure 7.13: Distribution of the regularization parameter, σ_r. The distribution is calculated by the use of kriging variances.

Figure 7.14 shows the evolution of the objective function versus the number of function evaluations. In 10-15 function evaluations the objective function is reduced by two orders of magnitude. The resulting permeability field is shown in Figure 7.15. The deterministic method has provided a match of the production data in few iterations. The price is that most of the geological consistency from the prior is lost in the result. The matched field clearly shows artifacts which are inconsistent with the geologic features in the reference and the prior. A higher degree of regularization may lead to a result with less artifacts. However, in this particular case application of a higher regularization parameter led to poor convergence and a match could not be obtained within a reasonable number of reservoir simulations.

7.4 Predictive capabilities

The main reason for history matching an oil reservoir is of course to use the resulting reservoir model for prediction of future production. It is common knowledge in reservoir engineering that integration of geologic knowledge into the reservoir model is important for the model's predictability. Therefore, it

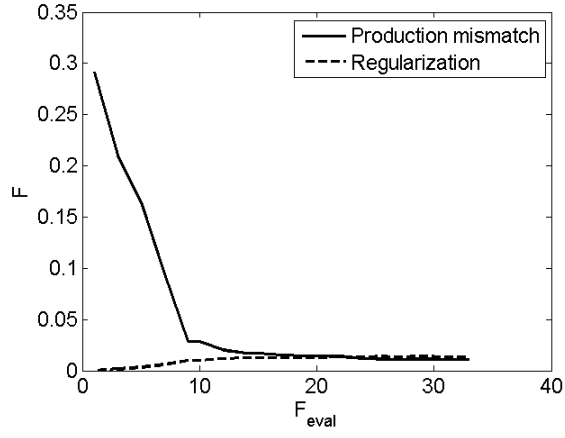


Figure 7.14: Evolution of the objective function during adjustment.

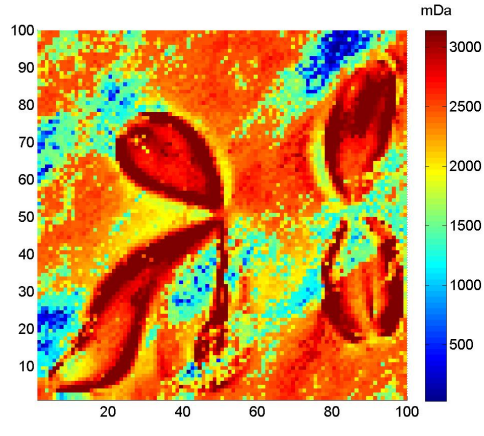


Figure 7.15: Resulting permeability after optimization.

is relevant to ask the question: Does the deterministic adjustment of the permeability deteriorate the predictable capabilities of the reservoir model? This question arises because the deterministic method does not conserve statistical properties such as the variogram. However, such geological properties are important pieces of information and are closely connected to the flow patterns in the reservoir.

To investigate the effects of the final adjustment on the predictability multiple history matches for a certain setup will be obtained. Again, gradual deformation will be applied first followed by an adjustment of the permeability with the Levenberg-Marquardt algorithm. The reference field is shown in Figure 7.16. The field is operated with one producer which is placed in the upper right corner of the field. Injectors are placed in the other corners, forming a symmetry element of the conventional quarter-nine spot configuration. History matching will be based on measurements of injector pressures and watercut at 8 times ending at 55 days. The production from 55 days up to 180 days can then be used to evaluate the predictability of the model.

Ten matches have been obtained from gradual deformation. In the gradual deformation ten outer loops have been allowed which means each of the obtained matches have been gradually deformed with up to ten realizations of permeability. Figures 7.17 and 7.19 show the ten matches of bottom hole pressures and watercut, respectively. Remember that only the measurements up to 55 days of production were matched. All the matches, except the pressure of injector 1, are centered around the reference. For injector 1 the matches are clustered at a slightly higher pressure than the reference.

The ten results from gradual deformation are now used in the deterministic methodology as priors. The results from adjusting the outcomes of gradual deformation with the Levenberg-Marquardt method are shown in Figures 7.18 and 7.20. It is evident from the figures that the adjustment has decreased the spread of the matches around the reference. And, more importantly, the unmatched production after 55 days is still described well. Off course, this result is specific to the setup considered here. As we saw earlier, a complex setup of wells can make history matching more challenging and geologic artifacts may arise in that connection. In such cases the predictability of an adjusted permeability field may be impeded. Thus, it would be mistaken to consider the results shown here as being general to all history matching problems. However, the fact that the predictions made with the adjusted fields are better indicates that the proposed methodology can be used with success in some cases.

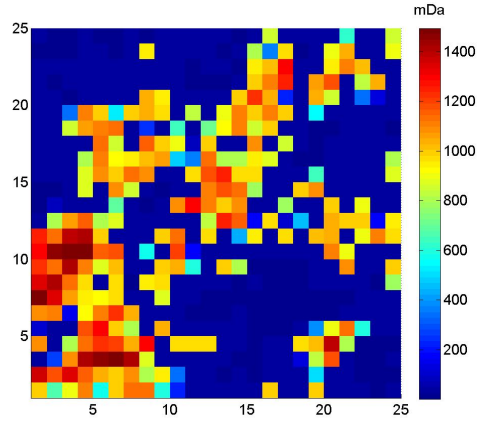
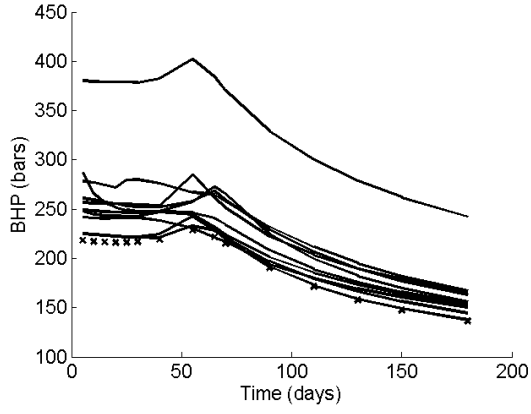
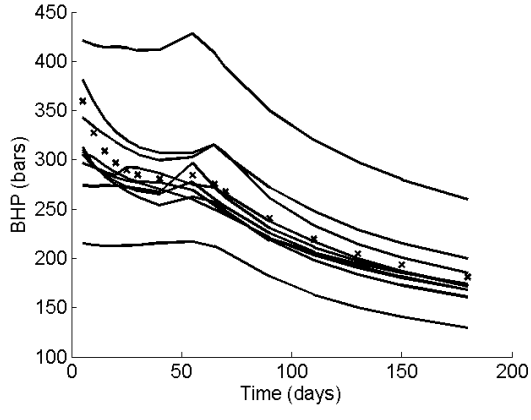


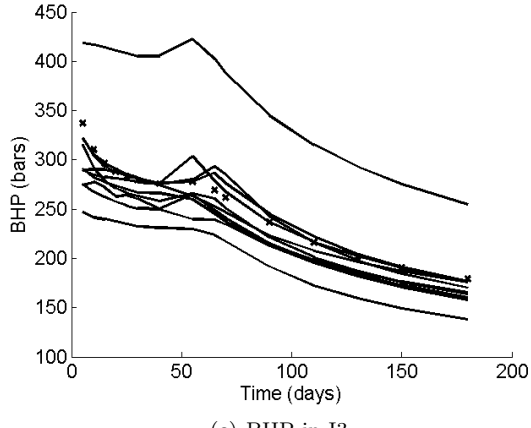
Figure 7.16: Reference field. A quarter nine spot well configuration is applied. One producer is placed in the upper left corner and injectors are placed in the other corners.



(a) BHP in I1.

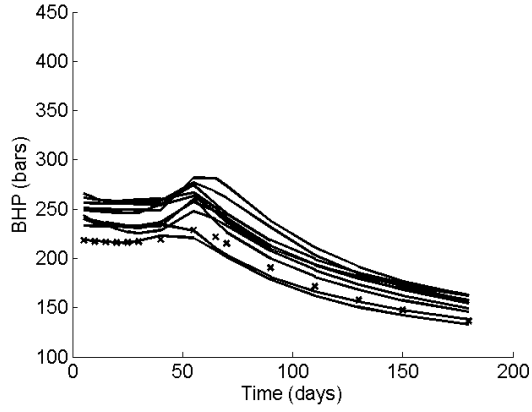


(b) BHP in I2.

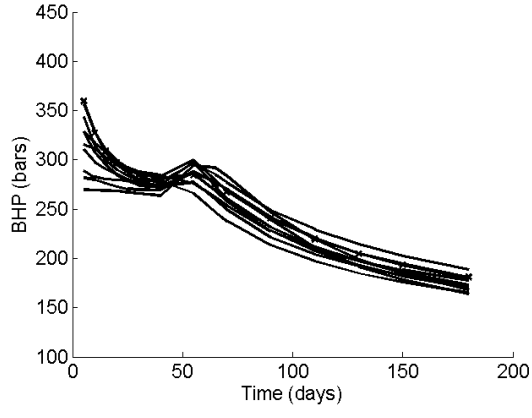


(c) BHP in I3.

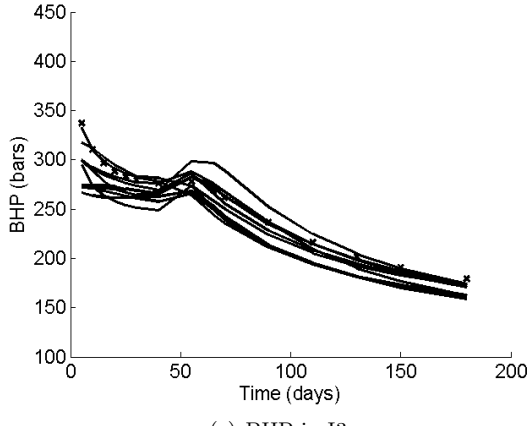
Figure 7.17: Matched injector pressures for ten cases only using gradual deformation. The reference data is indicated by ×. Only the first eight reference data (until 55 days) were used in the matching.



(a) BHP in I1.



(b) BHP in I2.



(c) BHP in I3.

Figure 7.18: Matched injector pressures for ten cases where the deterministic method has been used to adjust the match. The reference data is indicated by $\times$. Only the first eight reference data (until 55 days) were used in the matching.

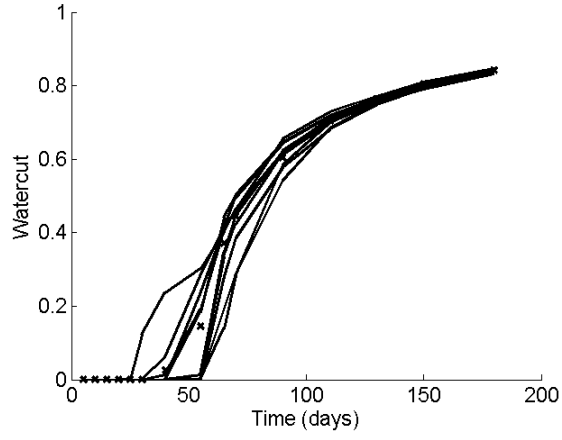


Figure 7.19: Matched watercuts for ten cases only using gradual deformation. The reference data is indicated by $\times$. Only the first eight reference data (until 55 days) were used in the matching.

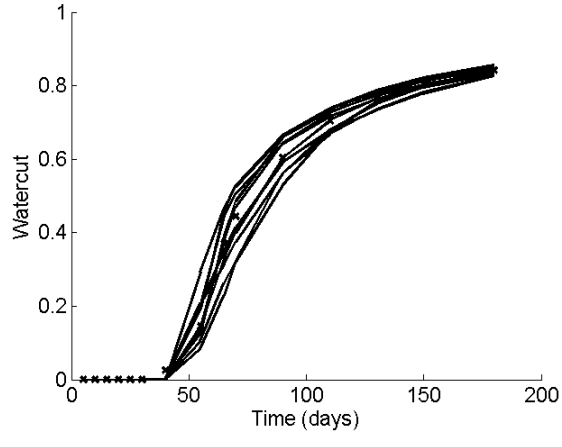


Figure 7.20: Matched watercuts for ten cases where the deterministic method has been used to adjust the match. The reference data is indicated by $\times$. Only the first eight reference data (until 55 days) were used in the matching.

7.5 Limitations of gradient-based methods

In most of the cases in the previous sections the gradient-based method performs well and is able to provide a good match of the production data within an acceptable number of reservoir simulations. However, the case considered in Section 7.3.1 illustrates some of the problems associated with gradient-based methods. With a too loose regularization the result may be dominated by artifacts which are inconsistent with prior geological knowledge. In this section the limitations of the gradient-based methods are investigated further.

In the following setups only permeability varies with location. Porosity and other parameters are held constant over the entire reservoir domain. The reservoir is defined on a $80 \times 80 \times 1$ block centered grid with grid blocks of the size $1m \times 1m \times 1m$. A facies model of the reference reservoir is shown in Figure 7.21(c). A simple cookie-cut [1] technique is used to populate each facies type such that there is some degree of heterogeneity within the two facies types. Two injectors inject water and production takes place with three producers. The specifications of the wells are given in Table 7.2. Water and oil are assumed incompressible.

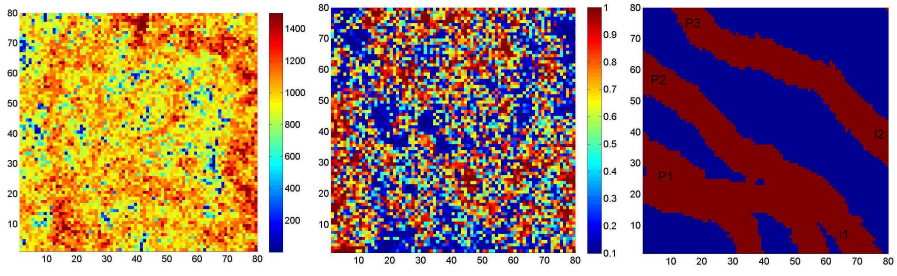
Table 7.2: Specifications of the wells.

Well	Type	i	j	Constraint
I1	Inj	68	5	Rate ($1m^3/day$)
I2	Inj	75	40	Rate ($1m^3/day$)
P1	Prod	10	25	BHP ($10bar$)
P2	Prod	10	55	BHP ($10bar$)
P3	Prod	20	70	BHP ($10bar$)

7.5.1 Homogeneous prior

The first case considered here uses a homogeneous prior of $432mDa$ which corresponds to the volume average of the reference field. To avoid overcorrection near the wells a spatially varying regularization parameter is applied. It is calculated by equation 7.13. Figure 7.22(b) shows the spatial variation of the regularization parameter. The Levenberg-Marquardt method is used to adjust permeability in order to achieve a better match of injector pressures and watercut. The matched data is depicted in Figure 7.23 together with the reference and initial data. The corresponding adjusted permeability field is shown in Figure 7.22(a). The evolution of the objective function is shown in Figure 7.24 which shows that the objective function is only reduced with approximately 50 pct.

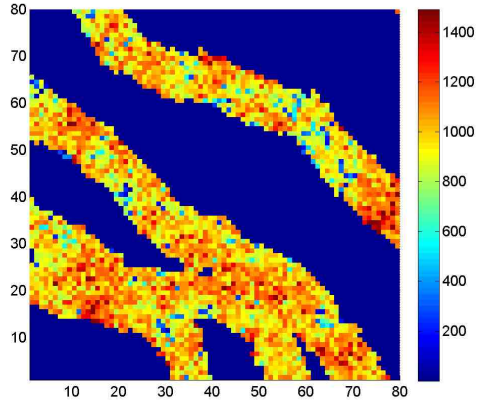
[1] Cookie-cut: The permeabilities within each binary facies are populated by pasting the permeabilities from two different realizations onto the facies model.



(a) Distribution for facies 1. $\mu = 994.9mDa$.

(b) Distribution for facies 2. $\mu = 0.52mDa$.

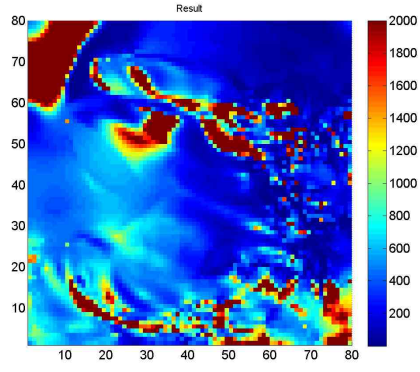
(c) Reference facies distribution.



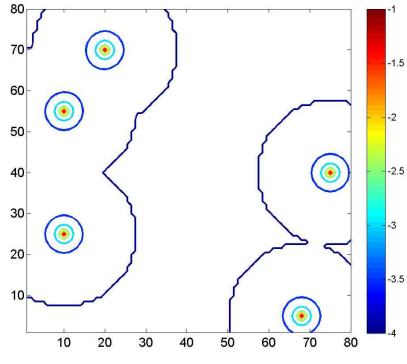
(d) Resulting reference. $\mu = 431.6mDa$.

Figure 7.21: On top: The permeability fields used to form the references case together with the facies model. Bottom: The resulting permeability field after a simple cookie-cut technique has been applied. The locations of the wells are indicated in the facies model.

after eight iterations in the Levenberg-Marquardt method. Further investigations with varying degrees of regularization and more iterations did not result in significant improvements of the match. The use of a gradient-based method on a binary field with large contrasts in permeability between each facies does not perform well. This is because the deterministic method operates with continuous variables and does not take the binary geology into account. Consequently, the gradient-based method is not able the construct high permeable channels but seeks to adjust the permeability in a continuous manner. The matched permeability field (Figure 7.22)does not look natural and is dominated by artifacts from the optimization.

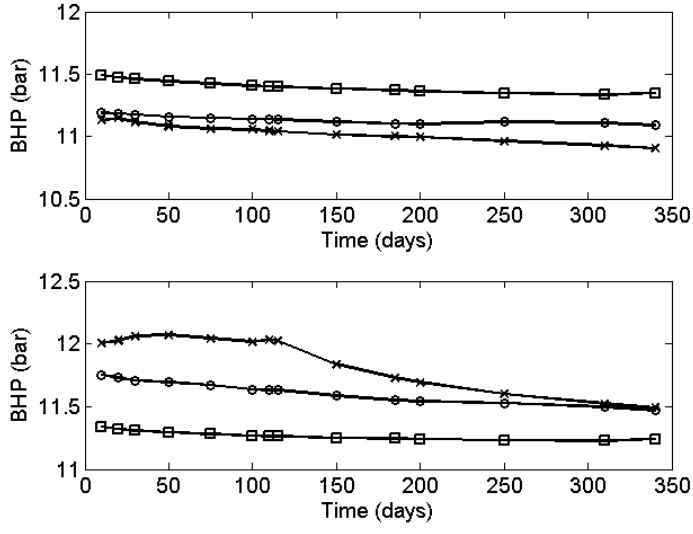


(a) Result.

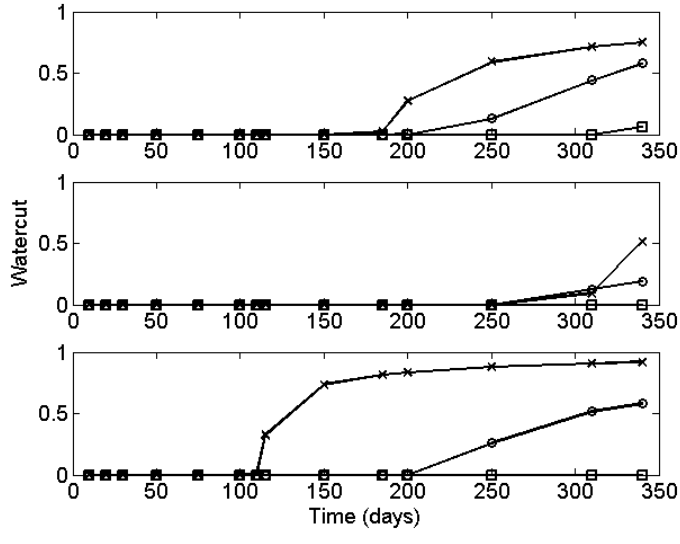


(b) $\log_{10} \sigma_r$.

Figure 7.22: Resulting permeability field. The results are from the case with a homogeneous prior and regularization calculated by the use of the kriging variance.



(a) Injector BHP.



(b) Watercut.

Figure 7.23: Production data. ×: Reference, ○: Match, □: Initial. Results are for the case with a homogeneous prior and regularization calculated by the use of the kriging variance.

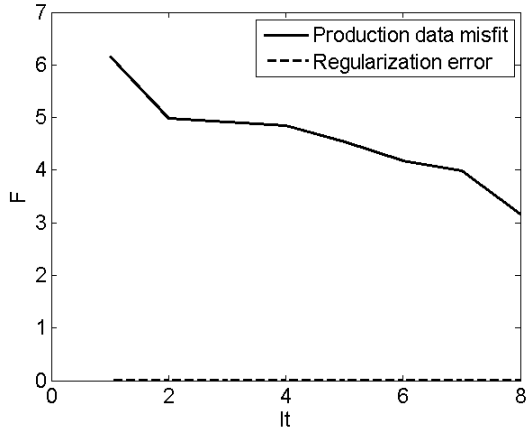


Figure 7.24: Evolution of the objective function in the case with with a homogeneous prior and regularization calculated by the use of the kriging variance.

7.5.2 Unregularized case with smooth initial guess

As illustrated above, optimization based on the Levenberg-Marquardt method was not able to match the production data when a homogeneous prior field was used in the regularization term. This is because the reference field is composed of channels with high permeability contrasts. The initial homogeneous field is a poor starting point for the optimization because it produces flow patterns significantly different from the reference. In this section a slightly modified field is used as initial guess for the optimization. In addition, regularization is not applied to the problem in order to increase the performance of the optimizer. The prior is now moderately heterogeneous with smoothly varying permeabilities around the wells. The principal direction of the heterogeneity is 45° which makes the initial guess mimic the channel structure from the reference. The areas between the wells are homogeneous with permeability $432mDa$ as before. The initial guess is shown in Figure 7.25(a). Again, the injector pressures and watercuts are matched. Figure 7.25(c) shows the evolution of the objective function which is reduced by four orders of magnitude within 20 function evaluations. The matched data is shown in Figure 7.26(a+b). The figure shows that the production data is matched very well. The matched permeability field is depicted in Figure 7.25(b). The Levenberg-Marquardt method has restored the high permeable channels between the injectors and the producers. However, extreme values have been introduced in certain areas. In a postprocessing step unlikely and extreme values may be filtered out but this has not been investigated here. Compared to the case where a homogeneous prior was used the match is significantly better. This is mainly because the moderate heterogeneity ensures that the initial guess results in better initial flow patterns. The

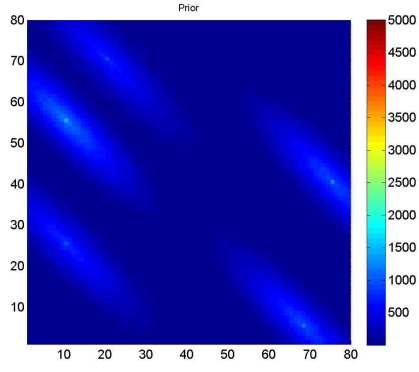
example shows that the choice of initial guess or prior is extremely important. The two initial fields used here result in very different matches; the homogenous prior leads to a bad match, whereas the moderately heterogeneous field leads to a perfect match. The example also illustrates the non-uniqueness of history matching problems. The field obtained with the moderately heterogeneous initial field leads to a perfect match of the measured production data but is different from the reference field.

7.5.3 Prior with wrong geometry

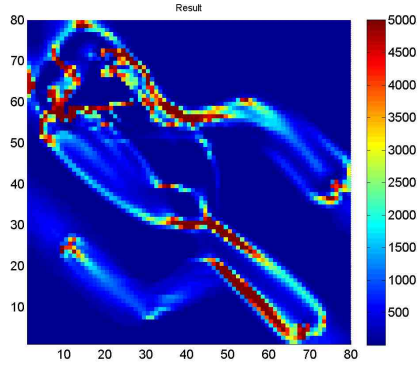
A more heterogeneous field is now used as prior for the regularization term. The prior is depicted in figure 7.27(a) and has a channel structure. The channel structure connects injector 2 to producer 2 which means that the fluid flow will be significantly different from the reference (see figure 7.21(d)). Now the prior is conceptually in accordance with the reference but the channels are not connected correctly. Figure 7.27(c) shows the evolution of the objective function. The objective function is reduced with less than an order of magnitude within 35 function evaluations. The matched production data is shown in Figure 7.28 which shows that the Levenberg-Marquardt method has improved the match but is still not able to match the production accurately. The resulting permeability field is given in Figure 7.27(b) which shows that the optimizer has increased the permeability between injector 2 and producer 3. This is done in order to compensate for the wrong connectivity in the prior. The optimizer has therefore tried to increase the flow to producer 3 in order to match the reference data. Again, the optimizer has introduced extreme values in order to change the flow paths in the reservoir.

7.6 Summary of deterministic methods

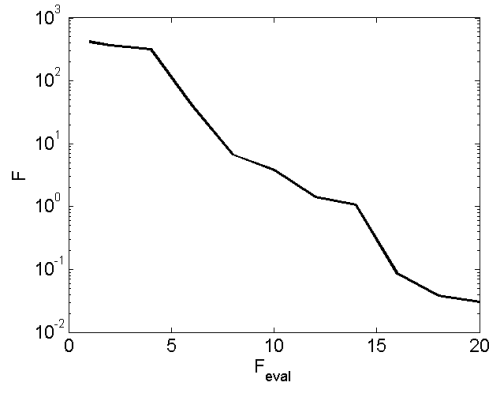
The examples presented in the current section illustrate some of the limitations of gradient-based methods. Such methods are generally challenged when the reservoir geology is complex and has features such as channels with a high contrast in permeability. Especially categorical variables are difficult to deal with when applying a gradient-based method and it may be very hard to honor geological knowledge when the field is history matched. Through regularization some handle on the geology is provided but a too severe regularization may lead to restrictive computation times because of slower convergence. However, for fields where the properties vary smoothly a gradient-based method can be the most efficient choice. A sequential strategy where a geostatistically based method is used to search for good initial guesses or priors for the gradient-based method may be a suitable approach. Such a method benefits of the efficiency of the gradient-based method but minimizes some of the related drawbacks such as the smoothing effects.



(a) Initial guess.

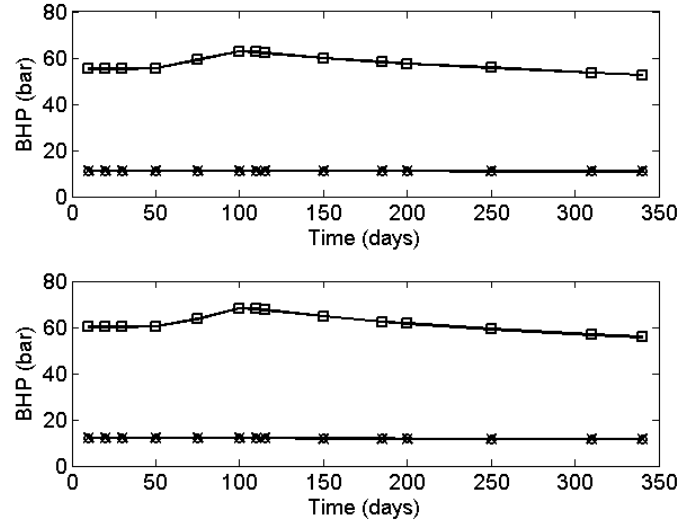


(b) Resulting field.

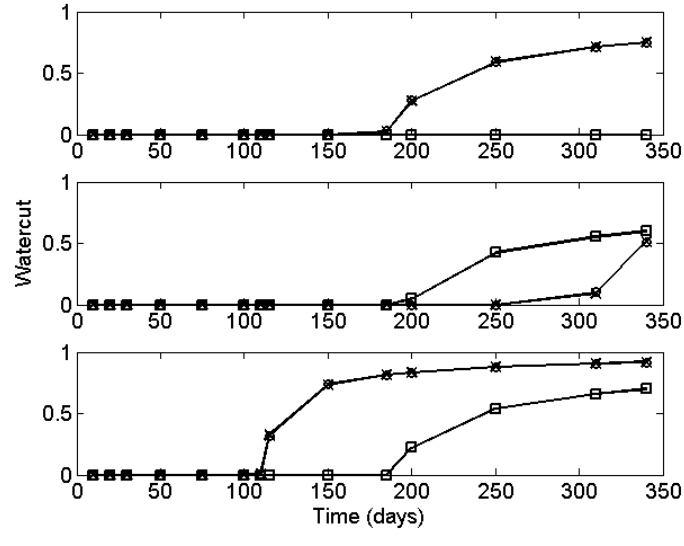


(c) Objective function.

Figure 7.25: Result from matching with an unregularized setup with a kriged field as the initial guess. The kriging is conditioned to hard data at the wells.

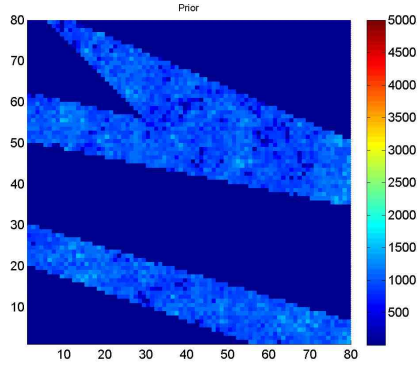


(a) BHP.

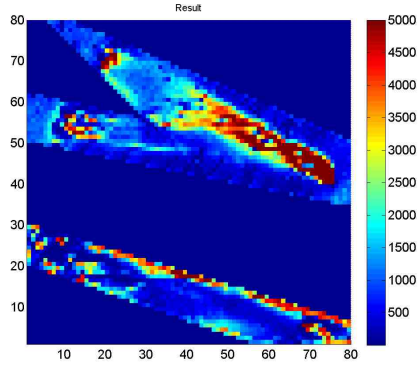


(b) Watercut.

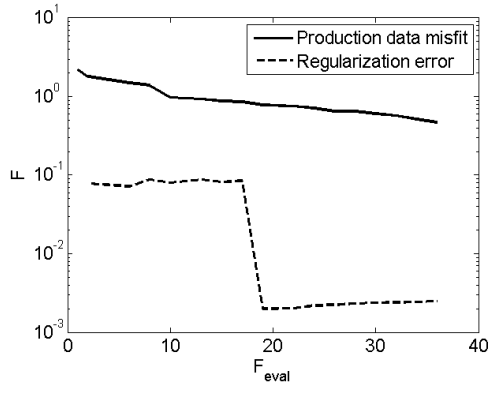
Figure 7.26: History matching result with the field shown in Figure 7.25(b) as initial guess. No regularization has been applied. ×: Reference, ○: Match, □: Initial. The mismatch of the pressures do not exceed 1 pct.



(a) Prior.

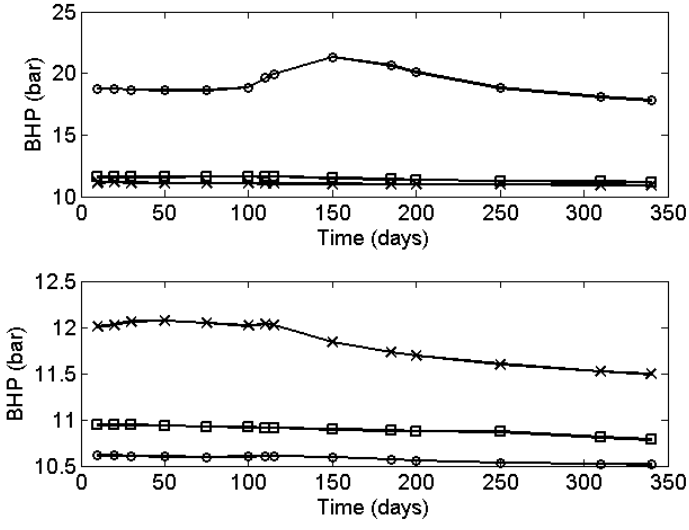


(b) Result.

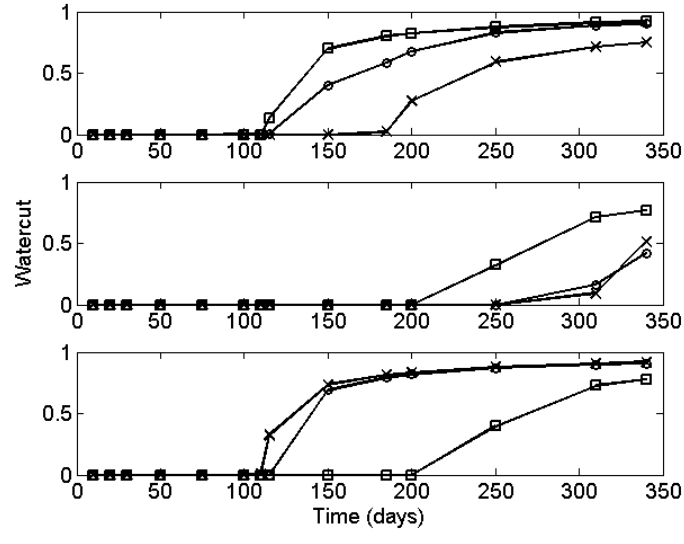


(c) Objective function.

Figure 7.27: Prior and result in the case with with a wrong facies distribution. There is connectivity between all producers and an injector but producer 2 is connected to the wrong injector. The jump in the objective function is the result of a reduction of the regularization parameter.



(a) Injector BHP.



(b) Watercut.

Figure 7.28: Production data in the case with with a wrong facies distribution. ×: Reference, ○: Match, □: Initial.

Chapter 8

Hybridization of PPM with gradient information

In the previous chapters two geostatistical methods for history matching were presented. In addition, deterministic methods were discussed and some of the limitations of the gradient-based methods were investigated. Until now the only combination of geostatistical methods with gradient-based methods has been the sequential methodology presented in Section 7.3 where gradual deformation is used to search for improved priors for regularization and initial guesses. The methodology is simple to implement because it is sequential. However, the sequential workflow also limits the efficiency because the two methods are separated from each other. This means that the problem of ensuring geological consistency of the reservoir model is not resolved. In this chapter a hybrid method is proposed. The proposed method takes advantage of the probabilistic framework of the PPM method which is used to integrate qualitative gradient information. In the following sections the background of the method is explained and some example uses are presented. In the final sections some more complicated cases are discussed. The hybrid method has been presented in the paper of Johansen, Caers & Suzuki (2007). Some of the examples shown in the dissertation are also used in that paper.

8.1 Background for the development of a hybrid method

The details of the probability perturbation method are given in Chapter 6. The following presentation of a hybrid method is based on the concepts introduced previously and should be viewed in that context.

In order to introduce the methodology used to extract the qualitative information from the gradient we will start by considering a simple but instructive example. Figure 8.1 shows a binary permeability field where the white color indicates a high permeable sand facies and black indicates low permeable mud. A water injector is placed in the lower right corner and producers are placed in the lower and upper left corners. Let the permeability field shown in Figure 8.2 be an initial guess of the true reservoir model.

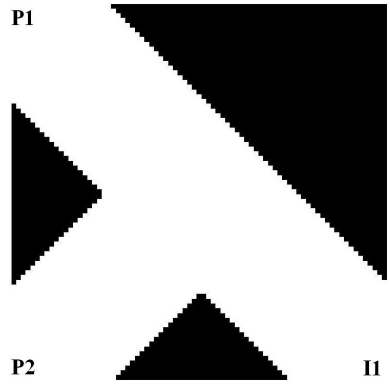


Figure 8.1: Binary permeability field used to illustrate the problems associated with the use of gradients in structural modeling. White color indicates high permeable sand facies and black indicates low permeable mud.

The initial guess is similar to the reference except that the producer in the lower left corner has been disconnected from the high permeable sand facies. Consequently, the production data from the reference and the initial guess are very different from each other because the area around the lower left producer will not be swept by the injected water. Figure 8.3 shows the sensitivity of the watercut at the upper producer at a particular time. The sensitivity of watercut in the lower producer is in this setup equal to zero at all times because this well never experiences a water breakthrough. The figure shows that the grid blocks forming the edges of the sand channel are the most sensitive ones. The sensitivity in the lower left area is very small compared to the sensitivity at the edges. Evidently, the sensitivity, and therefore also the gradient, does not suggest that the connectivity to the lower left producer should be re-established. A derivative-based optimization technique will therefore have

Figure 8.2: Prior field to illustrate the problems associated with the use of gradients in structural modeling. Compared to the field shown in Figure 8.1 the connectivity to the well in the lower left corner is not present.

difficulties establishing the connectivity and consequently, it will not be able to match the production data.

The following *squeezing function* is introduced:

$$\boldsymbol{g}_s = -sign(\boldsymbol{g}(\boldsymbol{x})) \odot (|\boldsymbol{g}(\boldsymbol{x})|)^n, \quad n \in]0;1], \tag{8.1}$$

where $\odot$ denotes an elementwise multiplication, $\boldsymbol{g}$ is the gradient, and $\boldsymbol{g}_s$ is a squeezed version of the gradient. The transformation makes small entries in $\boldsymbol{g}$ larger and large entries smaller.

If the sensitivity depicted in Figure 8.3 is subjected to the squeezing function it will take the form shown in Figure 8.5. The squeezed sensitivity field still indicates that the edges of the channel are influential on the watercut. However, compared to the unscaled sensitivity the squeezed sensitivity field shows a significant difference: the scaled sensitivity is high in a half-circular area placed in the lower left part of the field. This indicates that the sensitivity contains more information than what is immediately extracted from the unscaled sensitivity. This information is overshadowed by the higher sensitivities at the channel boundaries in the unscaled case. By scaling the sensitivity a valuable piece of *qualitative* information has been extracted: namely that the area in the lower left part of the field is more influential than other parts.

In the following discussion a methodology which makes use of such qualitative information to guide the PPM method is presented. The presentation is based on a binary representation of the reservoir geology. In this dissertation multi-categorical facies models are not treated. However, generalization to more complex cases involving more than two categories should be possible. Such a generalization could be the topic for future research within the area.

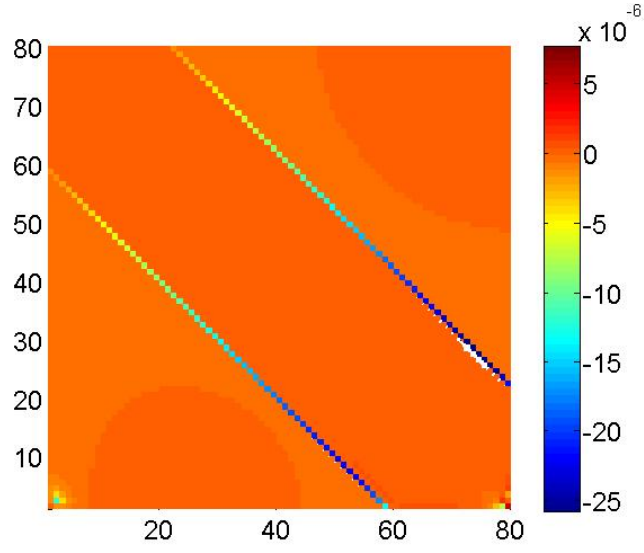


Figure 8.3: Sensitivity of watercut in P1 with respect to permeability at a particular time.

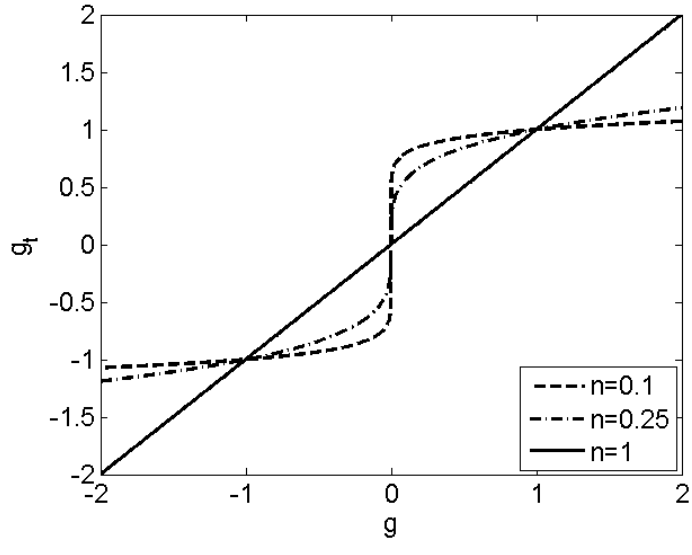


Figure 8.4: The squeezing function with two different exponents. The function makes small entries numerically larger and makes large entries numerically smaller.

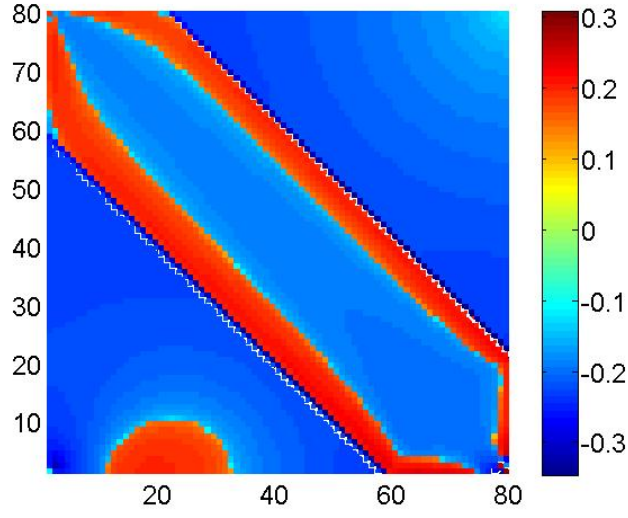


Figure 8.5: Sensitivity of watercut in P1 at a particular time. The figure is scaled by the use of the squeezing function (8.1) with $n = 0.1$.

Inclusion of the the qualitative gradient information into history matching requires a robust method to extract the information from the gradient. One way of extracting such information is to apply a truncation based scaling. A component in the gradient, G_i, can be scaled as:

$$G_{scl,i} = \begin{cases} -1 & \text{if}, G_i < -c \cdot \mu_{\boldsymbol{G}} \quad \text{(Add channel)} \\ 1 & \text{if}, G_i > c \cdot \mu_{\boldsymbol{G}} \quad \text{(Remove channel)} \\ 0 & \text{otherwise} \qquad\quad \text{(No change)} \end{cases}, \tag{8.2}$$

where $\mu_{\boldsymbol{G}}$ is given as

$$\mu_{\boldsymbol{G}} = \frac{1}{N_{blk}} \sum_{i=1}^{N_{blk}} |G_i| \tag{8.3}$$

and c is a number set by the user. $\mu_{\boldsymbol{G}}$ is the mean absolute value of the elements of the gradient.

The scaling given by equation (8.2) filters out the numerically smallest components of the gradient by setting the value of the scaled gradient equal to 0 in those grid blocks. If a component is smaller than $-c \cdot \mu_{\boldsymbol{G}}$ the scaled gradient is assigned the value -1 at that grid block. Likewise a component larger than $c \cdot \mu_{\boldsymbol{G}}$ is assigned the value 1. Recalling that the negative gradient direction is guaranteed to be a descent direction, a negative value of the gradient will be taken

as an indication that the permeability should be increased and a positive value implies a decrease, i.e. channel facies should be added or removed, respectively.

8.1.1 Integration of gradient information in PPM

The gradient information is calculated for a particular permeability field. This field should be the current best field, i.e. the field which has resulted in the lowest objective function value so far. Since the gradient suggests *changes* to the permeability field rather than suggesting an improved field directly, the gradient should always be related to the permeability field for which it was calculated.

Let $\boldsymbol{i}$ be a binary realization of indicator variables. Facies one will be attributed to the indicator 0 and facies two has the indicator 1. The scaled gradient $\boldsymbol{G}_{scl}$ is calculated for a certain history matching setup using the permeability field dictated by $\boldsymbol{i}$. The key idea in the proposed method is that a new binary facies indicator map given as:

$$\boldsymbol{G}' = \boldsymbol{i} - \boldsymbol{G}_{scl} \tag{8.4}$$

provides a good indication of how the permeability field might be improved.

As the framework for the PPM method is based on probabilities any additional information about the permeability field, e.g. seismic or the gradient information, needs to be formulated in terms of probability, too.

The qualitative gradient information can be formulated in terms of probabilities by the following expression

$$\boldsymbol{P}(A|D) = a\boldsymbol{G}'(A|D) + (1-a)P(A), \tag{8.5}$$

where $P(A|D)$ denotes the probability of the event A given the qualitative gradient information, D, a is a scalar which controls the impact of the gradient information. For $a = 0$ the conditional probability reduces to the marginal distribution of A. For $a = 1$ the impact of the gradient is largest. Consequently, a is referred to as the *degree of trust* in the gradient. The gradient information in the form of $P(A|D)$ needs to be combined with the conditional probability used in PPM, $P(A|C)$, where C denotes the production data mismatch. This can be done by using the τ-model which is introduced in Section 6.1 where it is used to combine the conditional probability $P(A|C)$ with the local probability function, $P(A|B)$. Inclusion of the gradient information into PPM necessitates the use of the τ-model twice, i.e.

1. Combination of gradient information with traditional PPM, i.e. the computation of $P(A|C,D)$ from $P(A|C)$ and $P(A|D)$

2. Combination of $P(A|C,D)$ and $P(A|B)$ during sequential simulation

Compared to traditional PPM the first step is new. The second use of the τ-model is performed when a new facies indicator is simulated at a grid note and is done in traditional PPM as well as in the proposed method.

The following algorithm to integrate qualitative gradient information is proposed:

1. Define a vector $\boldsymbol{G}'(A|D)$ as $\boldsymbol{i}^{(k)} + \boldsymbol{G}_{scl}$
   - $\boldsymbol{G}'(A|D)$ is formed by adding the current best realization, $\boldsymbol{i}^{(k)}$, to the scaled gradient. $\boldsymbol{G}'(A|D)$ indicates how the permeability can be improved. D denotes the qualitative gradient information.
2. Truncate $\boldsymbol{G}'(A|D)$ to the interval $[0; 1]$
   - The scaled gradient may suggest to remove channel facies at grid blocks which do not have channel facies associated or it may suggest to add channel facies to grid blocks which already have the channel facies associated. Therefore, a truncation is needed to ensure that the entries in $\boldsymbol{G}'(A|D)$ are either zeros or ones.
3. Define $\boldsymbol{P}(A|D) = a\boldsymbol{G}'(A|D) + (1-a)P(A)$
   - The parameter a is used to control the contribution of the gradient information to the PPM algorithm and is denoted *degree of trust* in the gradient.
4. Define $\boldsymbol{P}(A|C) = (1-r_c)\boldsymbol{i}^{(k)} + r_c P(A)$
   - This step is the traditional PPM step which forms $\boldsymbol{P}(A|C)$ which is adjusted to improve the match in an inner loop.
5. Combine $\boldsymbol{P}(A|C)$ and $\boldsymbol{P}(A|D)$ to form $\boldsymbol{P}(A|C, D)$ (τ-model)
   - The τ-model is used to combine the probability field from regular PPM with the gradient information.
6. Input $\boldsymbol{P}(A|C, D)$ as soft probability constraint to Snesim
   - The optimal r_c can now be found by application of for instance the Dekker-Brent algorithm. This is similar to traditional PPM. The only difference is that gradient information is now used to guide the PPM method.

As in traditional PPM a number of outer loops over the seed for Snesim are carried out until a sufficiently good match has been obtained or until a maximum number of iterations is made. Application of the gradient in every outer iteration is not a requirement. It may prove advantageous to apply the gradient information in every second or third iteration or so and then use traditional

PPM in the iterations in between. It is the user's task to device a suitable strategy for when to use the gradient information. A priori it is difficult to lay out such a strategy but experiences from previous uses may be helpful.

An instructive example which illustrates the idea behind the method is based on the reference permeability field depicted in Figure 8.6. The reference has a distinct channelized structure with high permeable channels oriented in the South-East direction. The permeability of the channel facies is $1150mD$ and the permeability of mud facies is $1mD$. A gridding of 80 by 80 has been used for the model. The field is operated by two injection wells and three producing wells. The exact locations of the wells are given in Table 8.1.

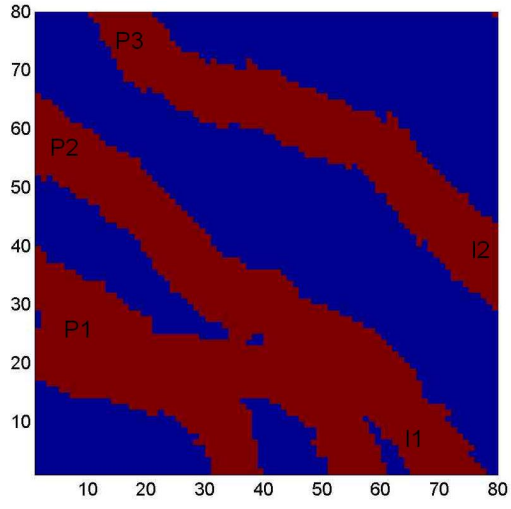


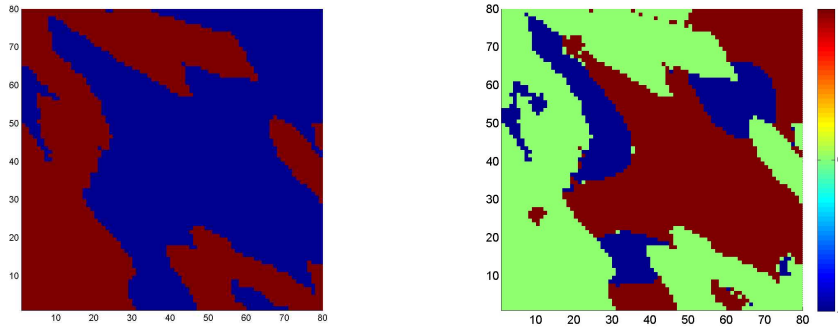
Figure 8.6: Reference facies distribution. The red color indicates high permeable sand facies and blue is low permeable mud. Injectors are placed in the right ends of the channels and production is taking place in three wells in the opposite side.

Table 8.1: Specifications of the wells. The wells are placed such that the water front will sweep along the channels.

Well	Type	i	j	Constraint
I1	Inj	68	5	Rate $(5m^3/day)$
I2	Inj	75	40	Rate $(5m^3/day)$
P1	Prod	10	25	BHP $(10bar)$
P2	Prod	10	55	BHP $(10bar)$
P3	Prod	20	70	BHP $(10bar)$

Only water and oil is assumed to be present. The fluids are treated as incompressible and relative permeability is modeled by a Corey-type power law. The permeability field shown in Figure 8.7(a) is constrained to facies data at

the well locations. Watercuts and oil rates at the three producers are to be matched. If the gradient of the sum-of-squares objective function is calculated and scaled by equation (8.2) the scaled gradient shown on Figure 8.7(b) will appear. The scaled gradient clearly suggests that the connectivity between the producers and injectors needs to be restored.



(a) Permeability field.

(b) Scaled gradient.

Figure 8.7: The left figure: Permeability field. To the right: Scaled gradient. Scaling is done according to equation (8.2) with $c = 1$.

Figure 8.8 shows how the combined probability $P(A|C, D)$ varies with the perturbation parameter, r_c. The effect of the scaled gradient is seen in the low permeable areas which are disconnecting the wells. Figure 8.9 shows the corresponding realizations from Snesim. From the figure it is seen that the connectivity is restored at $r_c = 0.5$. If reservoir simulations are carried out for each of the realizations the plot shown on Figure 8.10 can be constructed. The solid line represents the objective function versus r_c in traditional PPM where the gradient information is not applied. The dashed line represents the realizations shown in Figure 8.10 where the gradient information is applied. Inclusion of gradient information has lowered the value of the objective function significantly. Another observation is that the objective function for the case using the gradient information appears to be smoother. The fact that the current best realization is used to form the gradient-based probability $P(A|D)$ may have a regularizing effect on the objective function. Consequently, the gradient-guided PPM may be easier to work with in the inner iterations of the algorithm compared to traditional PPM because of the regularizing effect on the objective function.

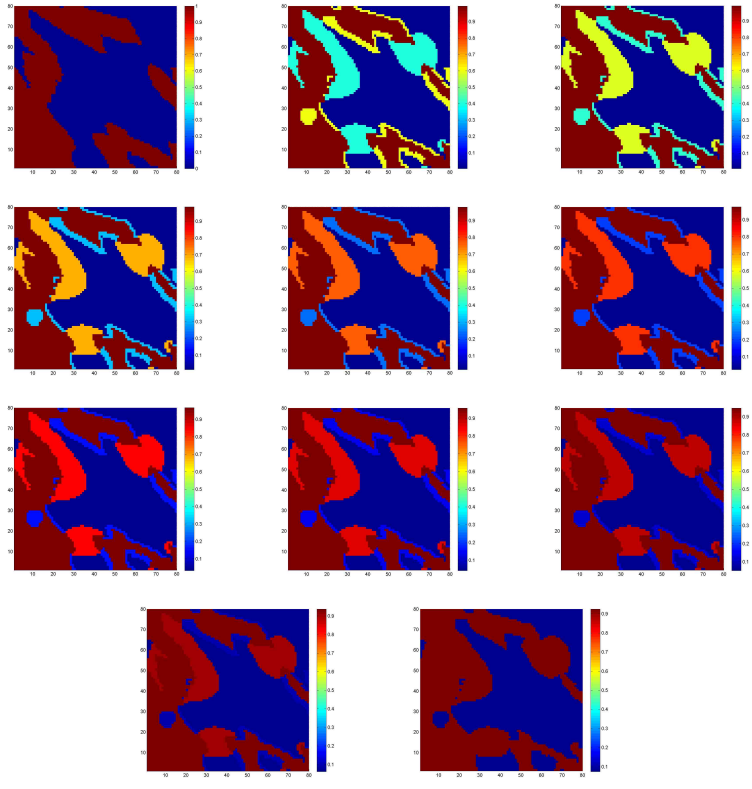


Figure 8.8: $P(A|C, D)$ for varying values of the parameter r_c. Starting with $r_c = 0.0$ in the upper left, r_c increases with 0.1 in the following figures. A degree of trust of 0.85 is used.

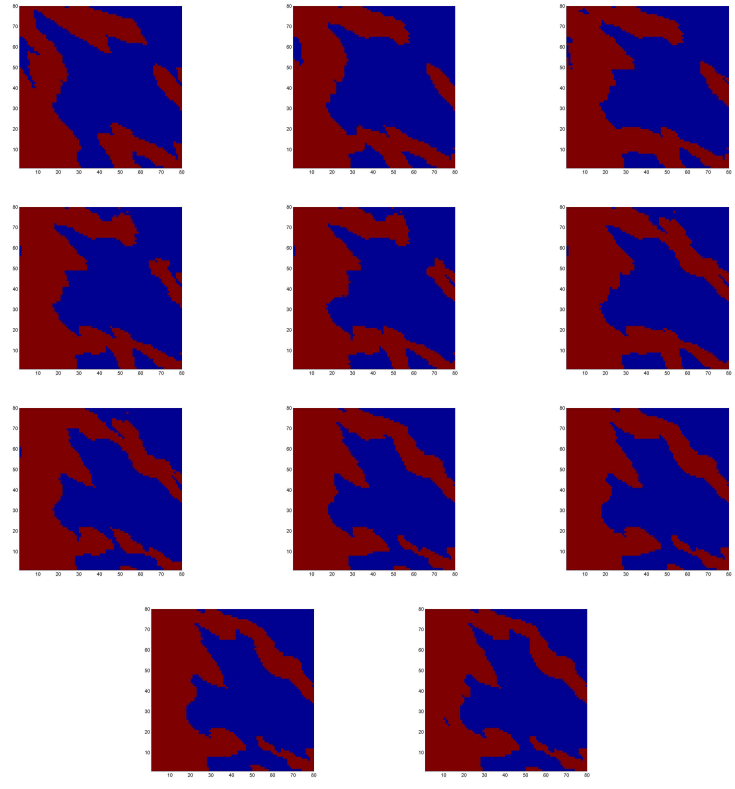


Figure 8.9: Resulting Snesim realizations for varying values of the parameter r_c. Starting with $r_c = 0.0$ in the upper left, r_c increases with 0.1 in the following figures.

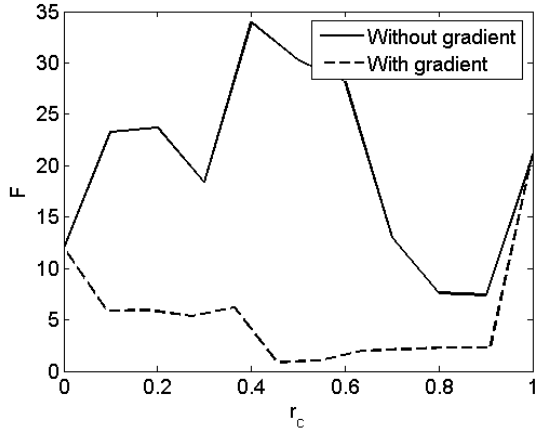


Figure 8.10: Objective function versus r_c. The dashed line corresponds to the steps shown in Figures 8.8 and 8.9 and the solid line is for traditional PPM without gradient information.

8.1.2 History matching with gradient guidance

The previous example was an instructive introduction to the concept of gradient-guided PPM. To evaluate the potential of the method a more elaborate test is performed. A full PPM history matching setup has been implemented with the option to apply gradient guidance at selected outer iteration steps. The framework for the case study is still the reference field shown in Figure 8.6 and the well configuration given in Table 8.1. The matched production data are watercut and oil rate. The gradient information is applied according to a simple strategy where gradient guidance is applied in every third outer iteration. In the other outer iterations traditional PPM is applied. Nine history matched realizations are constructed. Each of the matches is made with different starting guesses and different seeds during iterations. History matches from traditional PPM are also obtained and serve as comparison for gradient guided PPM and traditional PPM. A maximum number of ten outer iterations is allowed.

Figures 8.11 and 8.12 show the evolution of the objective functions during history matching for traditional PPM and gradient-guided PPM, respectively. If the objective functions are compared after about 35 function evaluations the gradient-guided PPM seems to perform better than traditional PPM. The gradient-guided PPM reaches a slightly lower end value for most of the objective functions. One of the objective functions in Figure 8.12 is not lowered much at the end of the history matching. The initial permeability field is in this case a poor starting point for the gradient guided method. However, it can be observed in Figure 8.11 that the same initial permeability field is not easily matched by the standard PPM method and a decrease is only obtained after a large number of function evaluations.

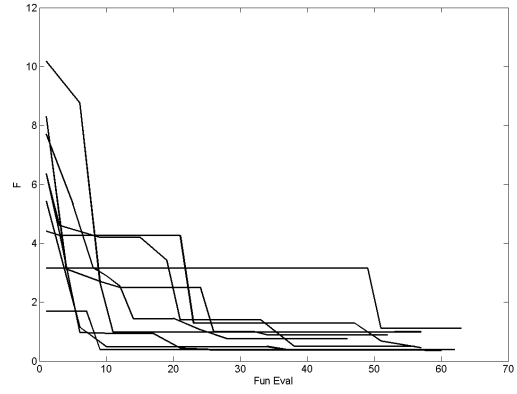


Figure 8.11: Evolution of the objective functions for nine history matches. Gradient guidance was not applied.

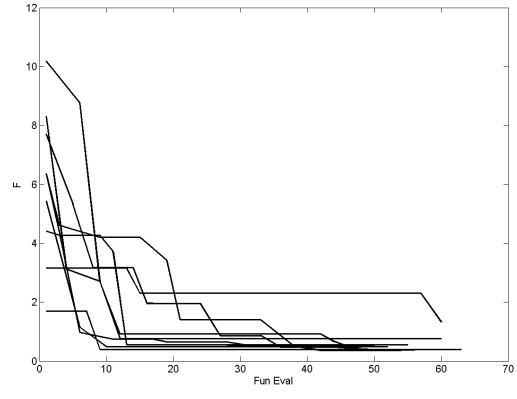


Figure 8.12: Evolution of the objective functions for nine history matches. Gradient guidance was applied in every third iteration.

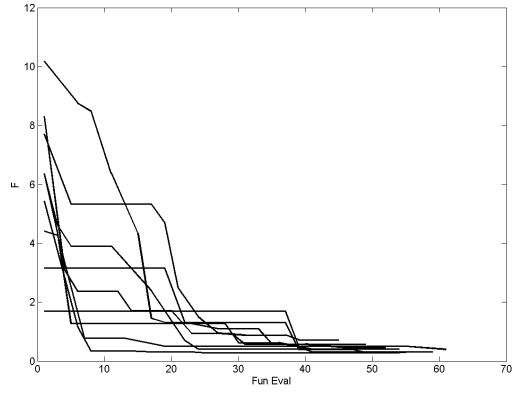


Figure 8.13: Evolution of the objective functions for nine history matches. Gradient guidance was applied by the use of equation (8.6).

An alternative way of integrating the gradient information has been investigated. The perturbation parameter, r_c, can be used to parameterize the gradient information more directly by defining an alternative probability:

$$P(A|D) = (1 - r_c)\boldsymbol{i} + r_c\boldsymbol{G}'(A|D) \tag{8.6}$$

where $\boldsymbol{G}'(A|D)$ is given by $\boldsymbol{i} + \boldsymbol{G}_{scl}$ and truncated to contain the indicators 0 or 1. The r_c parameter can now be used to perturb the probability from the current best field to the truncated field $\boldsymbol{G}'(A|D)$ which represents the current best field plus the qualitative gradient. Compared to the previously proposed algorithm the parameter a (degree of trust) is no longer needed. Now the trust in the gradient is integrated in r_c - if the gradient information leads to a decrease in the objective function the inner optimization loop will increase r_c. The same inner optimization as in traditional PPM can still be used since the adjustable parameter is still r_c and is limited to the interval $[0; 1]$. Since the current best field is now used to form the probability $P(A|D)$ more directly an even larger regularization effect on the objective function can be expected.

Figure 8.13 shows the nine objective functions for the case where equation (8.6) has been used to integrate gradient information. When the tails of the objective functions (from 30-50 function evaluations) are compared it is clear that the gradient-guided PPM in general reaches a lower end value of the objective function than in the case where gradient information is not applied. Secondly, the final value of the objective function is reached in fewer function evaluations. The individual matches are terminated after a varying number of function evaluations. This is because a termination criteria has been imposed on the inner loop. If r_c gets close to 0 the resulting realizations from Snesim will be identical. Therefore, the inner loop is terminated if r_c becomes smaller than a specified

threshold. Because of the regularizing effect from the inclusion of the gradient information it is likely that the Dekker-Brent algorithm reaches this threshold in fewer iterations. This may be part of the explanation for the improved performance of the gradient-guided PPM compared to traditional PPM. Figure 8.14-8.16 show the initial watercuts and matched watercuts from traditional PPM and gradient-guided PPM (using equation (8.6)). Both methods are able to match the watercut reasonably well. However, the matched watercut from the gradient-guided PPM appears to be more centered around the reference than is the case with traditional PPM.

The performance of the three history matching approaches is summarized in Tables 8.2 and 8.3. The number of function evaluations required to reach an objective function value of 0.5 or less is tabulated in Table 8.2. The improvement of convergence is also observed in the table. The gradient-guided methods generally reach an objective function value of 0.5 or less faster than traditional PPM. Table 8.3 shows the final values of the objective function after history matching. The final value is generally lower when gradient information is applied. Especially, the use of equation (8.6) to include the gradient information results in lower final values of the objective function.

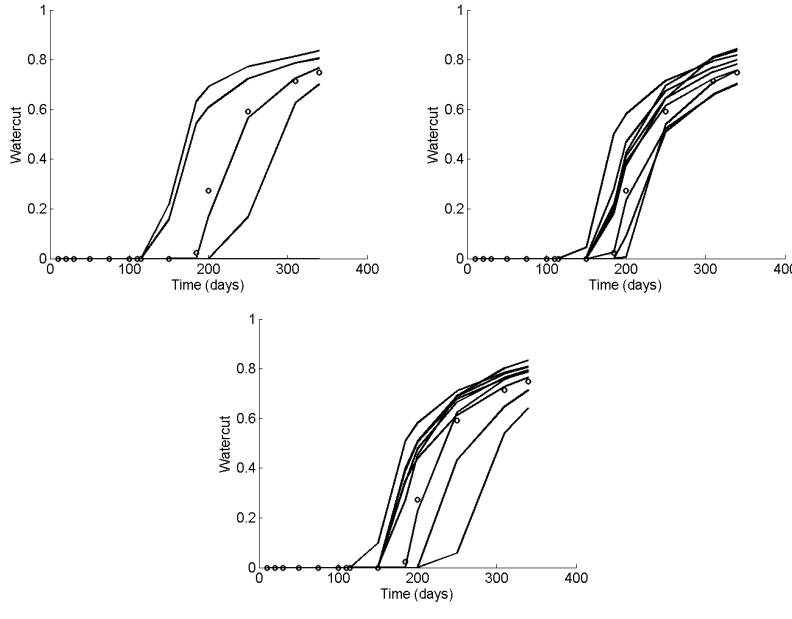


Figure 8.14: Nine matches of watercut for the three producers. The figures show watercuts from the unmatched initial fields. The reference is marked with ∘.

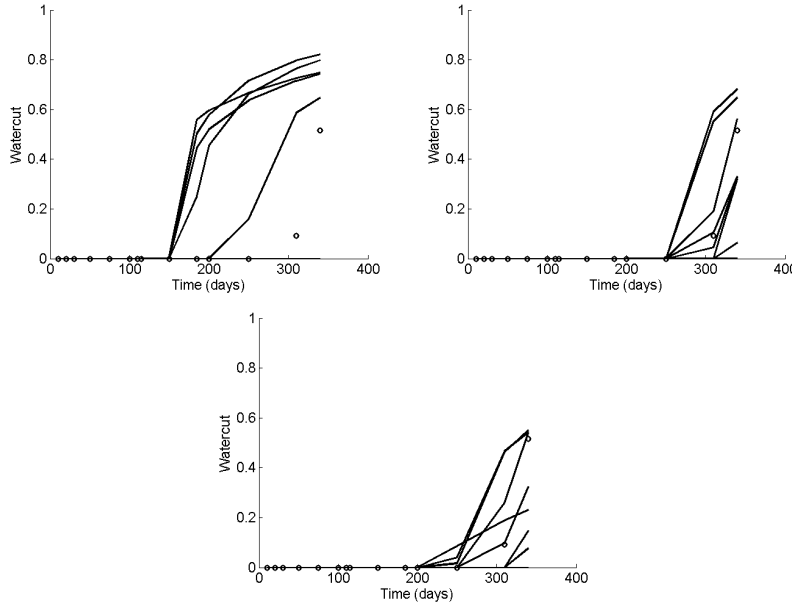


Figure 8.15: Nine matches of watercut for the three producers. The figures show the results from traditional PPM. The reference is marked with ∘.

Table 8.2: Number of function evaluations needed to reach an objective function value of 0.5. $> xx$ denotes that a value of 0.5 was not reached in xx function evaluations. gPPM^1 denotes gradient-guided PPM by the use of the algorithm outlined on page 127 and gPPM^2 implies that equation (8.6) was used to integrate gradient information. Bold face indicates the lowest number of iterations in each case.

Case	1	2	3	4	5	6	7	8	9
PPM	> 63	**9**	10	> 46	> 57	> 55	> 52	**21**	57
gPPM^1	> 60	**9**	10	**36**	46	42	> 60	32	> 55
gPPM^2	**39**	41	**8**	> 45	**39**	**24**	> 49	46	**19**

Table 8.3: Final values of the objective function. gPPM^1 denotes gradient-guided PPM by the use of the algorithm outlined on page 127 and gPPM^2 implies that equation (8.6) was used to integrate gradient information. Bold face indicates the lowest value in each case.

Case	1	2	3	4	5	6	7	8	9	Avg.
PPM	1.11	0.40	0.36	0.77	0.99	0.51	0.89	**0.38**	0.45	0.65
gPPM^1	1.32	0.40	0.38	**0.47**	0.49	**0.36**	0.77	0.50	0.56	0.58
gPPM^2	**0.29**	**0.31**	**0.27**	0.72	**0.39**	0.41	**0.56**	0.46	**0.39**	**0.42**

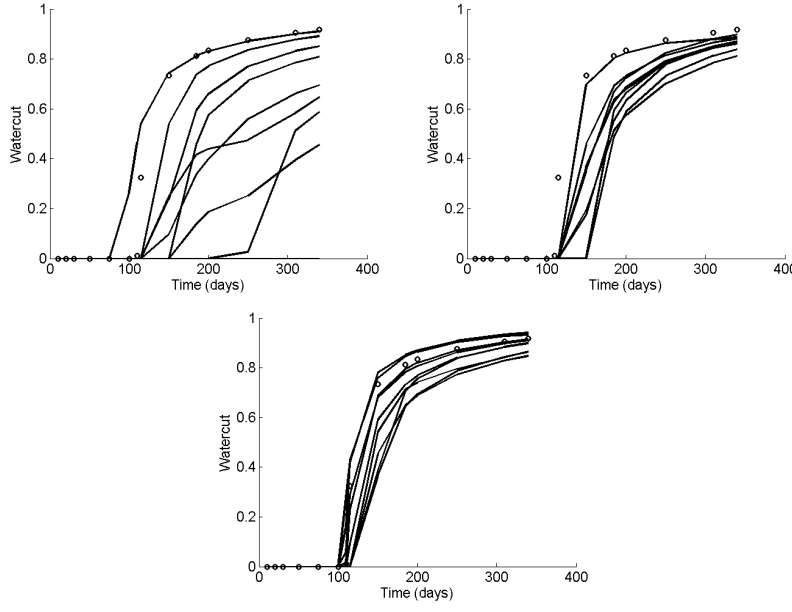


Figure 8.16: Nine matches of watercut for the three producers. The figures show the results from gradient-guided PPM (using equation (8.6)). The reference is marked with ∘.

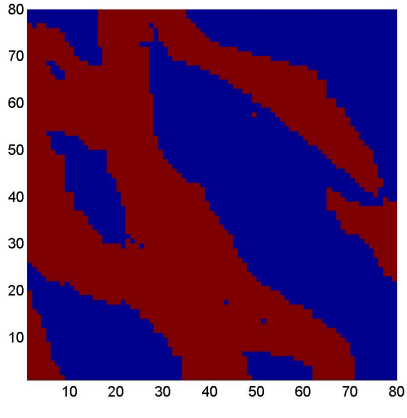


Figure 8.17: The realization associated with the lowest objective function value. The result is from a gradient-guided PPM history matching.

8.2 Applications on history matching cases

8.2.1 3D case with horizontal wells

The first case which is presented is based on a simple field which is represented by a binary facies model in three dimensions. The reference field is depicted in Figure 8.18. The field is composed of high permeable sand channels separated by a low permeable mud facies. The permeability of the sand is $1000mDa$ and the mud has a permeability of $1mDa$. The reservoir measures $100m \times 100m \times 15m$ and is operated by a horizontal injector placed in one side of the reservoir. Production takes place in a parallel producer in the other side. The fluid flow takes place in the direction of the channels. The field is discretized into a $50 \times 50 \times 5$ grid, yielding a 12500 gridblock model. The reservoir fluids are treated as incompressible. Consequently, it is not necessary to match both oil rate and water cut because only one producer is present and the water is being injected at a constant rate. The volumetric outflow is thus constant and the oil rate can be derived from the watercut. Therefore, history matching is limited to injector pressure and watercut. Production data up to 230 days is matched and the production from 230 up to 500 days is used to evaluate the predictive capabilities of the match. Figure 8.2.1 shows the matched pressure and watercut. The figure shows that the injector pressure is harder to match than watercut. There is a bias of the matched pressure which tends to be higher than the reference. The matched watercut is more evenly distributed around the reference. The spread of the predicted pressure and watercut increases with time. Figures 8.2.1 shows approximated 95 pct. confidence intervals for the predicted pressure and watercut. The confidence intervals are based on 15 matched fields. The field which led to the lowest objective function value is shown in figure 8.21 and the E-type of the 15 best matches is shown in figure 8.22.

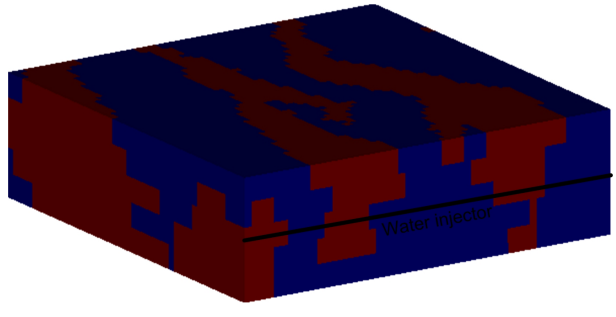
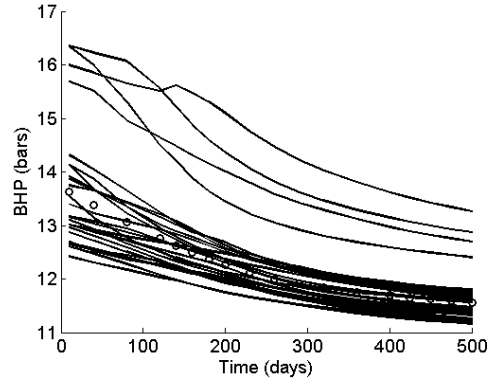


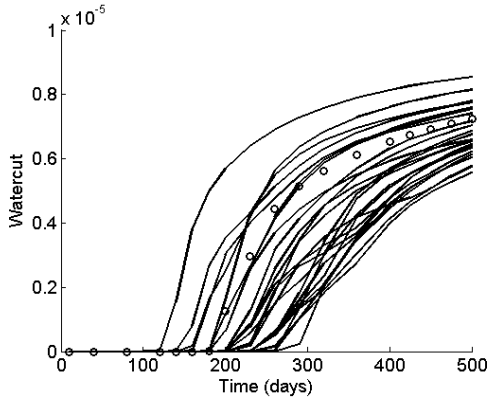
Figure 8.18: Reference permeability field. Blue color indicates low permeable mud and red indicates high permeable sand. Production is taking place in horizontal well parallel to the injection well. The wells are perforated in the entire reservoir interval. The field is discretized into a $50 \times 50 \times 5$ grid.

Table 8.4: Specifications of the wells. The wells are horizontal and drilled in the y-direction.

Well	Type	dir	i_{start}	i_{end}	j	k	Constraint
I1	Inj	y	1	50	1	3	Rate ($15m^3/day$)
P1	Prod	y	1	50	50	3	BHP ($10bar$)



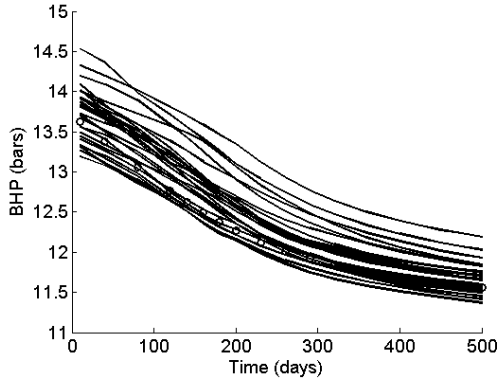
(a) Injector pressure.



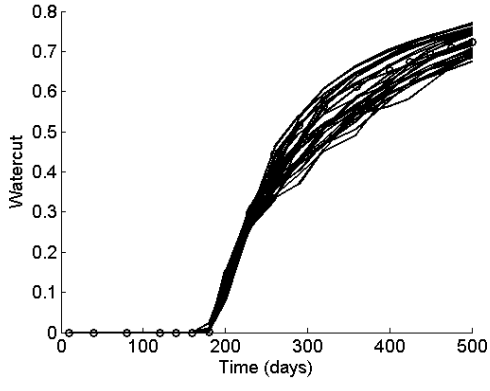
(b) Watercut.

Figure 8.19: Initial production data from the fields used to initiate history matching. The reference data is marked with circles.

14 matches have been obtained by traditional PPM without gradient guidance such that the effect from the gradient guidance can be evaluated. The production data for the fields used to initialize history matching is shown in Figure 8.2.1. The figure shows that the initial models generally lead to poor



(a) Injector pressure.



(b) Watercut.

Figure 8.20: Matched production data. The first nine measurements are used to match the production. The production from 230 days is not used to history match the reservoir model and can be regarded as projections of future production, i.e. forecasts.

matches. The matched production data for 14 history matches obtained with traditional PPM can be seen in Figure 8.2.1. The spread of the matched production data appears to be a larger compared to the matched production data shown in Figure 8.2.1 where gradient guidance was applied. The permeability field which led to the best match is shown in Figure 8.25.

Table 8.5 summarizes the number of function evaluations needed to terminate history matching for traditional PPM and gradient guided PPM. In seven cases the gradient guided method requires less function evaluations than traditional PPM. In three cases traditional PPM requires less evaluations than

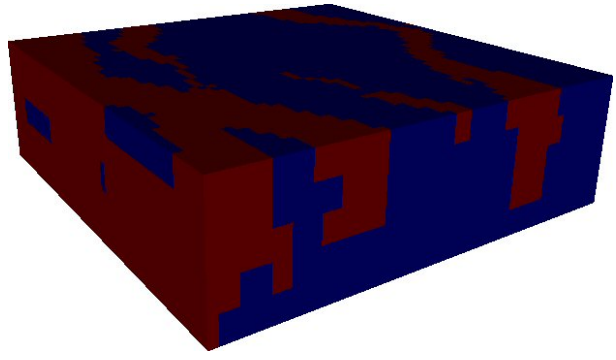


Figure 8.21: The matched permeability field which led to the best match of the production data (injector pressure and watercut).

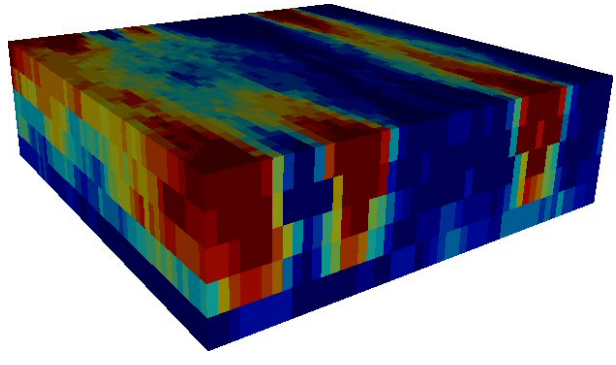
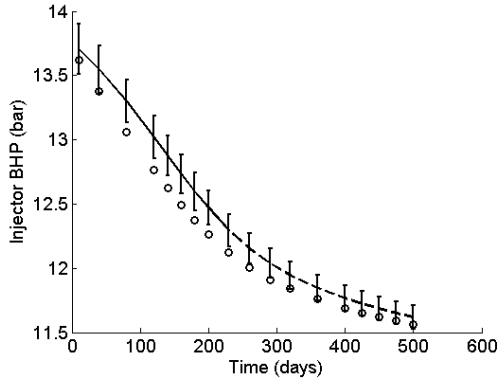
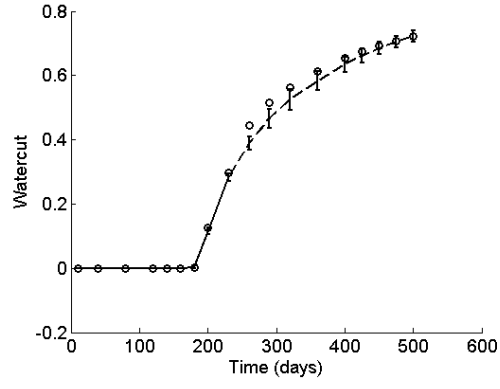


Figure 8.22: E-type of the 15 fields which resulted in an objective function less than 0.05.



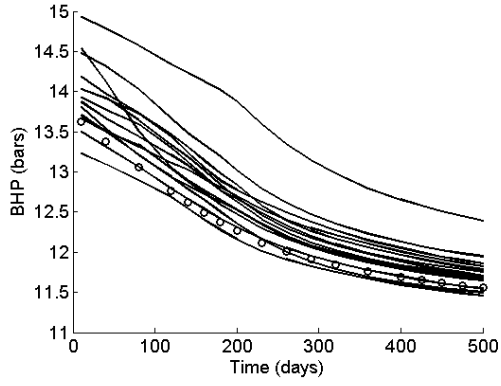
(a) Injector pressure.



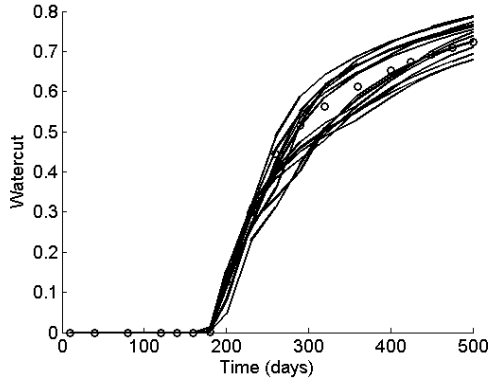
(b) Watercut.

Figure 8.23: Matched production data and predictions with 95% confidence intervals. The first 9 measurements are used to match the production. The production from 230 days is not used to history match the reservoir model and can be regarded as projections of future production, i.e. forecasts.

gradient guided method. The average number of function evaluations for 14 history matches are 34 for traditional PPM and 27 for gradient guided PPM. The results indicate that the gradient guided method in most cases performs better than traditional PPM. Table 8.6 shows the final values of the objective functions for the 14 cases with and without gradient guidance. In this case the gradient guided PPM does not perform as good as traditional PPM. However, traditional PPM leads to a very bad match in case number 3 where the gradient guided method performs significantly better.



(a) Injector pressure.



(b) Watercut.

Figure 8.24: Matched production data. The first nine measurements are used to match the production. The production from 230 days is not used to history match the reservoir model and can be regarded as projections of future production, i.e. forecasts. The results are from traditional PPM without gradient guidance.

Table 8.5: Number of function evaluations needed before termination of gradient guided PPM and traditional PPM. In seven cases the gradient guided method terminates after fewer function evaluations than traditional PPM. The average number of iterations for traditional PPM is 34 and 27 for gradient guided PPM.

case	1	2	3	4	5	6	7	8	9	10	11	12	13	14
PPM	28	36	49	50	48	4	40	49	32	5	49	23	20	36
gPPM	17	21	20	50	34	4	49	11	31	5	49	41	9	42

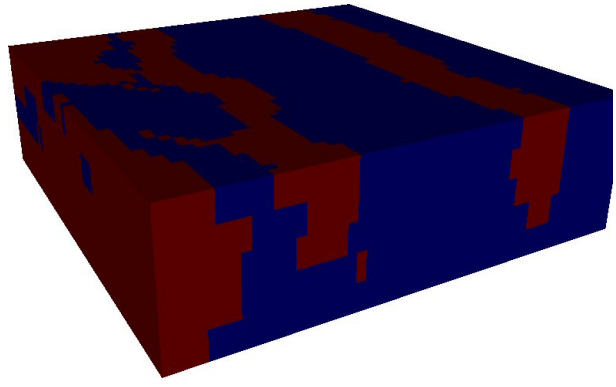


Figure 8.25: The matched permeability field which led to the best match of the production data (injector pressure and watercut). The result is from traditional PPM without gradient guidance.

Table 8.6: Final values of the objective functions with and without gradient information. The average values are 1.05 for traditional PPM and 0.72 for gradient-guided PPM.

case	1	2	3	4	5	6	7
PPM	0.10	0.14	5.77	0.22	1.00	0.14	0.43
gPPM	0.29	0.88	0.98	0.22	0.55	0.14	0.85
case	8	9	10	11	12	13	14
PPM	1.14	0.63	0.55	1.59	0.64	0.18	0.36
gPPM	0.90	0.86	0.55	1.59	0.94	0.76	0.24

8.2.2 3D case with horizontal and vertical wells

A more complicated field is now discussed. A reference field is shown in Figure 8.26. The reference field is generated by Snesim and is discretized into a $50 \times 50 \times 5$ grid and has the dimensions $100m \times 100m \times 15m$. Two facies are present: high-permeable channel facies ($1000mDa$) and low permeable mud ($1mDa$). Two injection wells are used to inject water and production takes place in three wells. The locations of the wells are given in Table 8.7 and are also indicated on the reference field. The wells are placed such that the injected water will sweep the high-permeable channels. A mix of horizontal and vertical wells is used and the wells are perforated in all grid blocks containing a well. Sequential simulations are conditioned to facies data along the wells. Water and oil are treated as incompressible. The production from the reference field is simulated until 500 days. Only the first 200 days of production will be matched and the remaining production data will be used to evaluate the predictive capabilities of the models. Nine matches have been obtained using traditional PPM and gradient-guided PPM with gradient guidance in every third outer iteration. A maximum of ten outer iterations is allowed. The production data associated with the nine initial reservoir models used to initialize history matching witch are shown in Figures 8.27, 8.28, and 8.29. The figures show that the initial fields lead to poor matches of injector pressure, watercut, and oil rate.

The matched production data for gradient guided PPM are shown in Figures 8.30, 8.31, and 8.32. The figures should be compared to the matches obtained with traditional PPM which are shown in Figures 8.34, 8.35, and 8.36.

Figures 8.30 and 8.34 show the matched injector pressures for gradient-guided PPM and traditional PPM, respectively. The gradient-guided PPM gives significantly better matches of the pressure in injector 1. The difference in the matches of pressure in injector 2 is not as marked but still shows that the gradient-guided method performs better. The same conclusion can be drawn from the matches of watercut and oil rates where the gradient-guided method performs slightly better than traditional PPM. The oil production rate from producer 2 is difficult to match. This producer is placed in a thin high-permeable channel which makes it difficult to match its production data.

Table 8.8 shows the final values of the objective function for traditional PPM and gradient guided PPM for the nine matches. In seven cases the gradient-guided method ends with a lower value than traditional PPM. In the two remaining cases the methods perform equally well. The mean of the final values are 1.34 and 0.87 for traditional PPM and gradient-guided PPM, respectively. The table thus confirms that the gradient-guided PPM method performs best in this history matching setup. Table 8.9 shows that the number of function evaluations is approximately the same for traditional PPM and gradient guided PPM. The field which led to the best match is depicted in Figure 8.33.

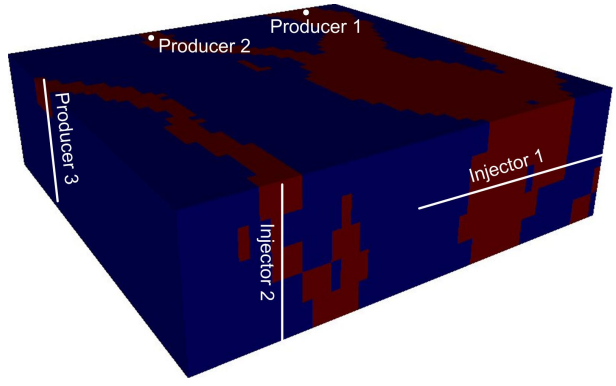


Figure 8.26: Reference permeability field. Blue color indicates low permeable mud and red indicates high permeable sand. Production is taking place in vertical wells and injection takes place in a vertical and a horizontal well. The field is discretized into a $50 \times 50 \times 5$ grid.

Table 8.7: Specifications of the wells. The wells are horizontal and drilled in the y-direction.

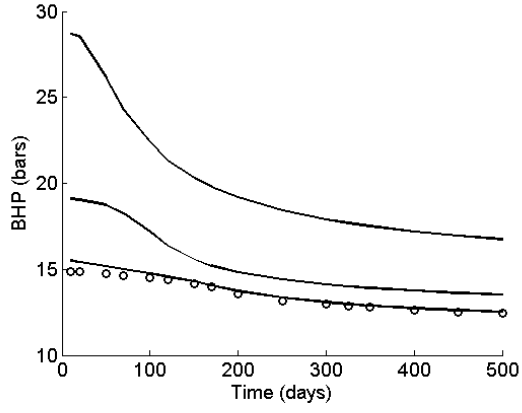
Well	Type	dir	i_{start}	i_{end}	j	k	Constraint
I1	Inj	y	1	25	1	3	Rate $(15m^3/day)$
I2	Inj	z	1	5	1	42	Rate $(5m^3/day)$
P1	Prod	z	1	5	50	9	BHP $(10bar)$
P2	Prod	z	1	5	50	33	BHP $(10bar)$
P3	Prod	z	1	5	36	50	BHP $(10bar)$

Table 8.8: Final values of the objective function. Bold face indicates the lowest value in each case.

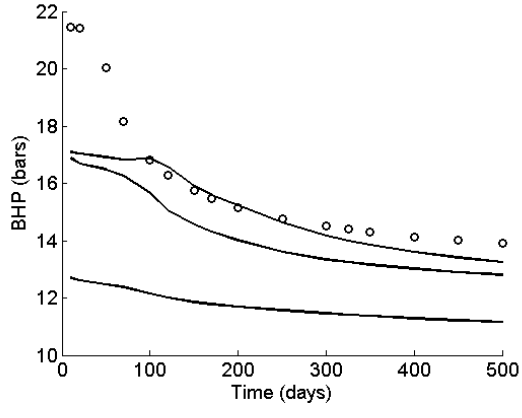
Case	1	2	3	4	5	6	7	8	9	Avg.
PPM	1.41	0.37	0.66	2.91	1.61	0.96	1.48	1.08	1.58	1.34
gPPM	**0.73**	**0.38**	0.66	**0.92**	**1.01**	**0.59**	1.48	**0.82**	**1.26**	**0.87**

Table 8.9: Number of function evaluations. Bold face indicates the lowest number in each case.

Case	1	2	3	4	5	6	7	8	9	Avg.
PPM	**43**	48	53	**45**	49	49	53	56	49	49
gPPM	49	**46**	53	49	**44**	**46**	53	**50**	**46**	**48**

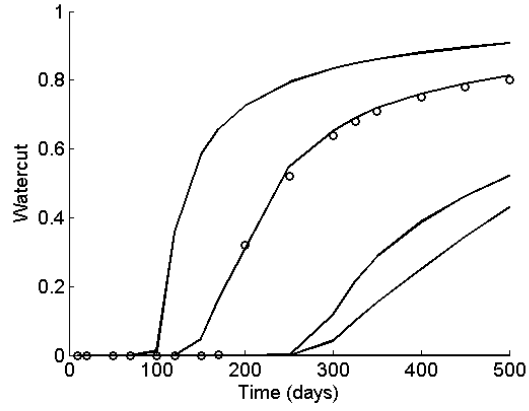


(a) Injector pressure 1.

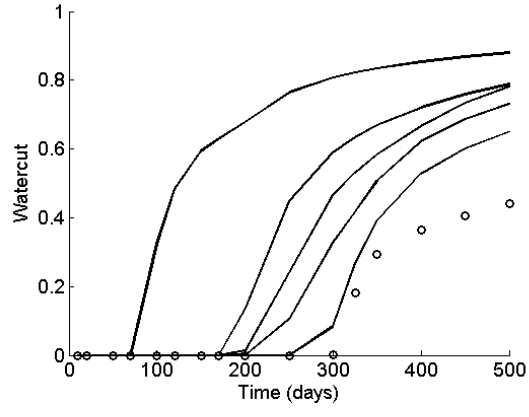


(b) Injector pressure 2.

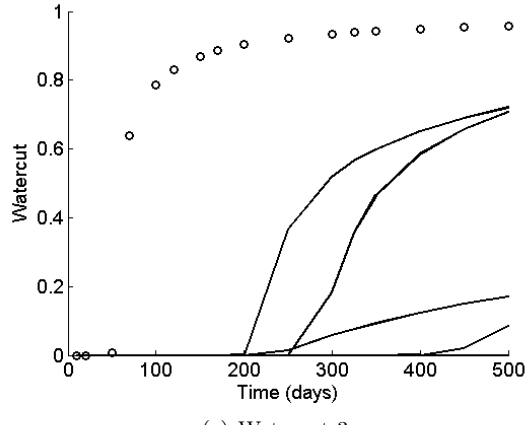
Figure 8.27: Initial production data data from the fields used to initiate history matching. The reference data is marked with circles. Some pressures have been omitted because they are too high for the axis scaling.



(a) Watercut 1.

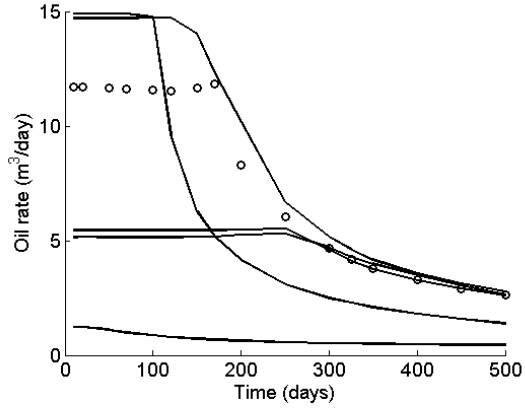


(b) Watercut 2.

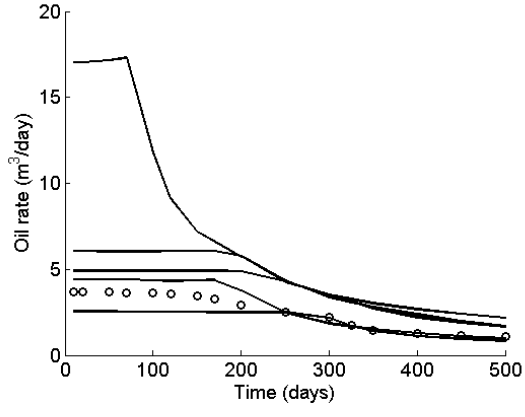


(c) Watercut 3.

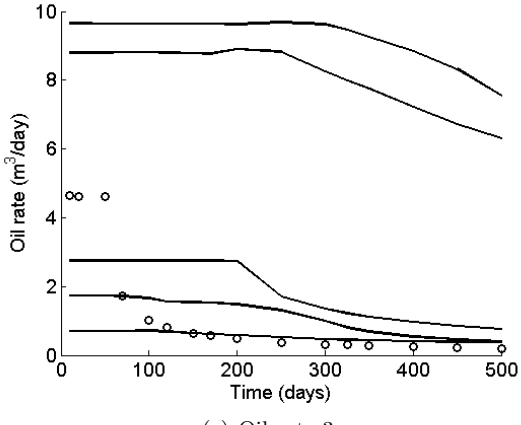
Figure 8.28: Initial production data data from the fields used to initiate history matching. The reference data is marked with circles.



(a) Oil rate 1.

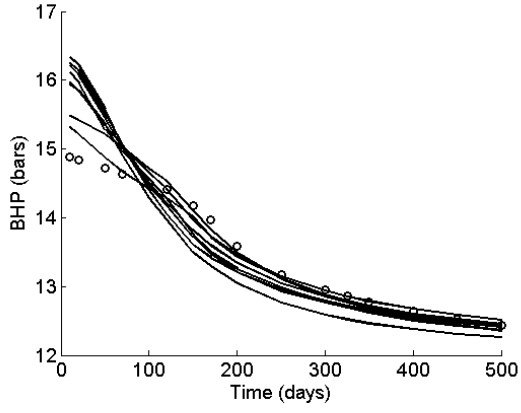


(b) Oil rate 2.

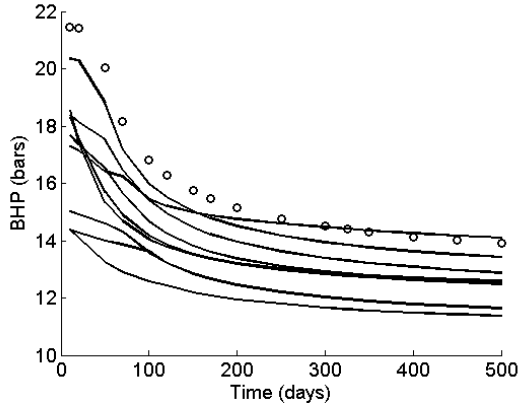


(c) Oil rate 3.

Figure 8.29: Initial production data data from the fields used to initiate history matching. The reference data is marked with circles.

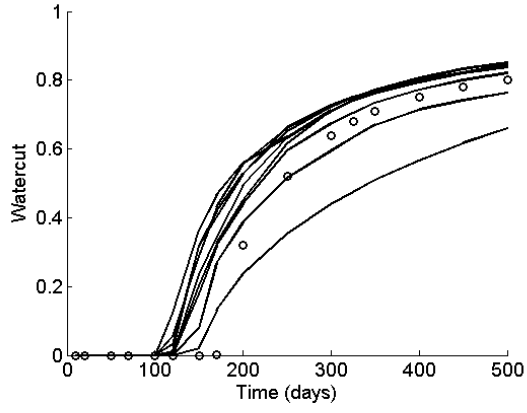


(a) Injector pressure 1.

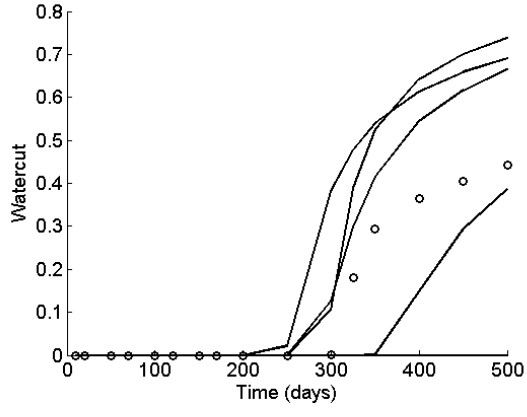


(b) Injector pressure 2.

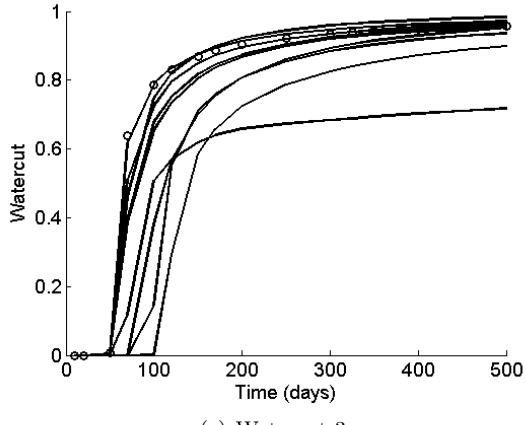
Figure 8.30: Matched production data. The reference data is marked with circles. The pressures from one match has been omitted because they are too high for the axis scaling. Only production up to 200 days has been matched. The data from 200 to 500 days is forecasted.



(a) Watercut 1.

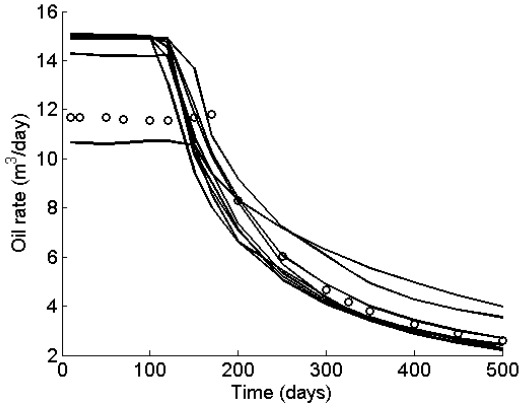


(b) Watercut 2.

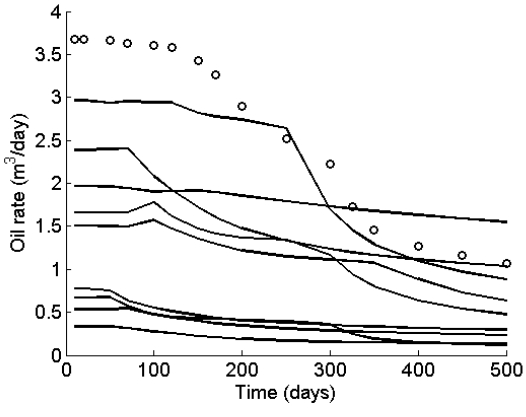


(c) Watercut 3.

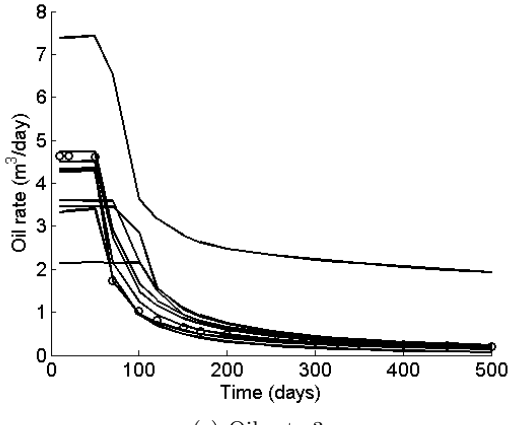
Figure 8.31: Matched production data. The reference data is marked with circles. Only production up to 200 days has been matched. The data from 200 to 500 days is forecasted.



(a) Oil rate 1.



(b) Oil rate 2.



(c) Oil rate 3.

Figure 8.32: Matched production data. The reference data is marked with circles. Only production up to 200 days has been matched. The data from 200 to 500 days is forecasted.

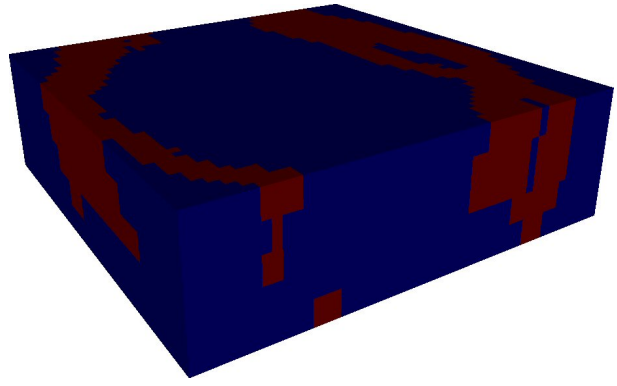
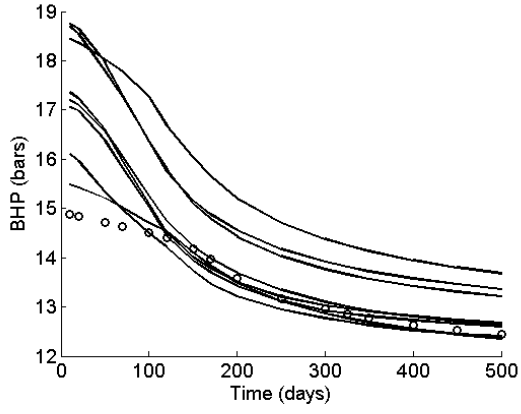
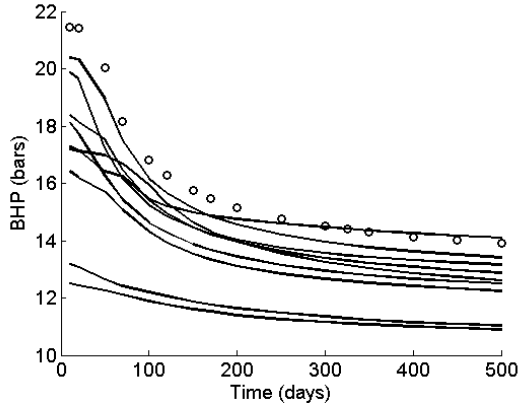


Figure 8.33: The matched permeability field which led to the best match of the production data. The result is from PPM with gradient guidance.

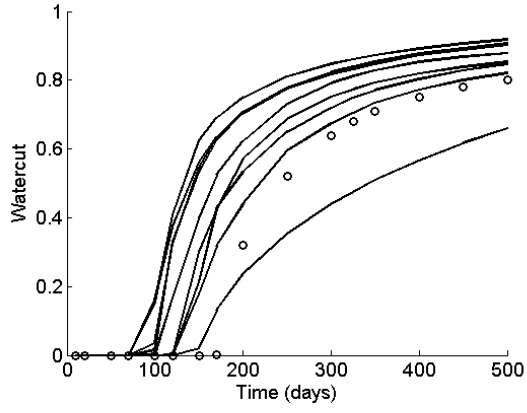


(a) Injector pressure 1.

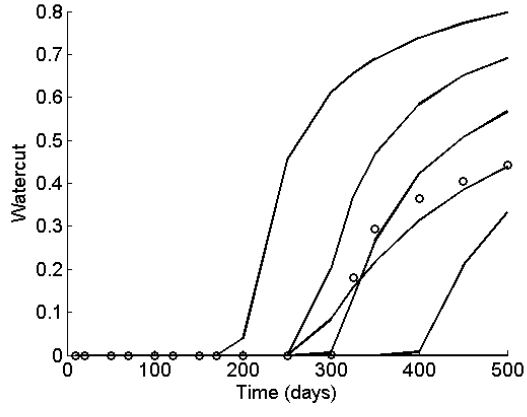


(b) Injector pressure 2.

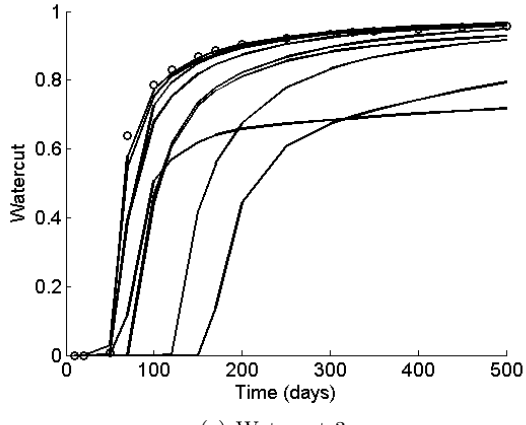
Figure 8.34: Matched production data. The reference data is marked with circles. The pressures from one match has been omitted because they are too high for the axis scaling. Only production up to 200 days has been matched. The data from 200 to 500 days is forecasted. No gradient guidance has been applied.



(a) Watercut 1.

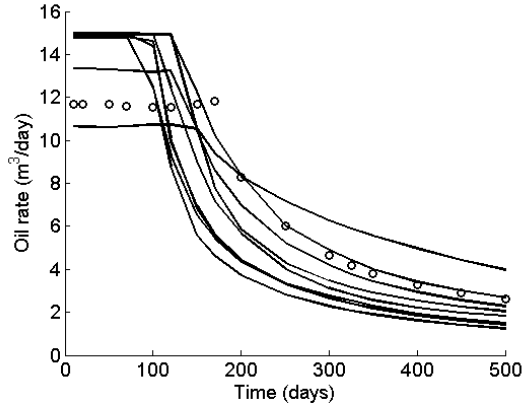


(b) Watercut 2.

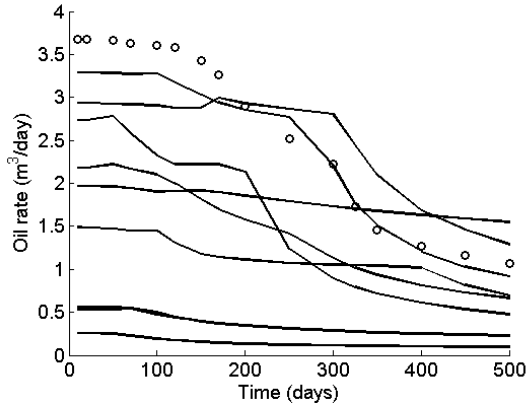


(c) Watercut 3.

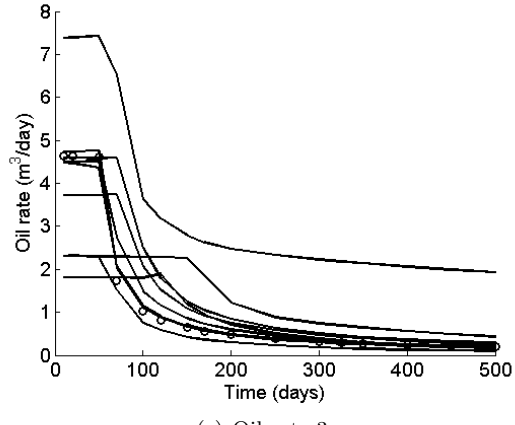
Figure 8.35: Matched production data. The reference data is marked with circles. Only production up to 200 days has been matched. The data from 200 to 500 days is forecasted. No gradient guidance has been applied.



(a) Oil rate 1.



(b) Oil rate 2.



(c) Oil rate 3.

Figure 8.36: Matched production data. The reference data is marked with circles. Only production up to 200 days has been matched. The data from 200 to 500 days is forecasted. No gradient guidance has been applied.

8.3 Summary of the proposed hybrid method

A hybrid method based on the framework of the probability perturbation method (PPM) has been proposed. This method includes qualitative gradient information into traditional PPM by the use of the τ-model proposed by Journel (2002). By filtering out numerically small elements of the gradient of the objective function it is possible to extract valuable qualitative information which can be used to guide the PPM method. The proposed gradient-guided PPM method has been applied to history matching examples in 2D as well as in 3D. In all cases the method performs better than traditional PPM, yielding lower end values of the objective function. The gradient-guided method reaches the final value in fewer iterations. This may be due to regularization effects as well as the qualitative information extracted from the gradient. An existing framework for traditional PPM can easily be extended to include the gradient guidance. This, however, requires that an efficient calculation of sensitivities of production data with respect to reservoir parameters is available. In the current setup an efficient adjoint method is applied to calculate these sensitivities.

Chapter 9

Conclusion

The PhD study has dealt with application and development of history matching procedures in a geostatistical framework. The main contribution from the project has been the development of a hybrid method which combines the geostatistical probability perturbation method (PPM) with qualitative gradient information. The hybrid method has been applied to history matching problems with two-dimensional and three-dimensional reservoir models. A general observation from these applications it that the inclusion of gradient information improves the performance of the PPM method. The main characteristics of the proposed hybrid method are:

- A scaling of the gradient of the objective function is performed in order to extract qualitative gradient information.
- Two ways of integrating the gradient information are presented. One of these make use of a parameter, denoted *degree of trust*, which controls the impact of the gradient information. If this parameter is set to zero the hybrid method reduces to traditional PPM.
- A simple strategy where gradient information is applied en every third outer iteration of PPM is applied.
- An improvement of the performance has been observed when gradient information is included in PPM.
- Inclusion of the gradient information has a regularizing effect on the objective function which appears to be more smooth with respect to changes in

the perturbation parameter. This results in better performance compared to traditional PPM.

Apart from presenting the hybrid method the thesis also contains an elaborate discussion on various history matching techniques from the literature. This includes:

- The probability perturbation method
- Gradual deformation
- Deterministic methods

The probability perturbation method is introduced in some detail as it makes up the basis for the main contribution of the work. It has been possible to extend PPM to include gradient information because of its probabilistic framework. This framework makes PPM flexible in the sense that many diverse information sources can be integrated into the method.

The gradual deformation method is a geostatistical parametrization which enables an automated history matching procedure to perturb Gaussian distributions in a systematic way. If the reservoir geology can be represented by a Gaussian distribution the gradual deformation method can be applied to history matching of production data. The method conserves second-order statistics such as variograms. The literature contains numerous works which take their basis in this method. PPM method appears to be more flexible than the gradual deformation method regarding the reservoir geology. However, in cases where the geology can be represented by a Gaussian distribution the gradual deformation has advantages, e.g. it allows for multiple deformation parameters which can increase the flexibility but also the complexity of the parameterization. The cases considered in this thesis have been based on Gaussian fields of continuous variables. However, the gradual deformation method has been extended to object-based models in published literature.

The deterministic optimization method Levenberg-Marquardt has been applied to history matching problems. In order to stabilize the history matching problem a Tikhonov regularization has been applied. Apart from improving the numerical stability of the problem the regularization also provides a way of ensuring some geological consistency of the result. The main conclusions regarding deterministic methods are:

- Efficient when the geological model exhibits smooth variations of the adjusted parameter.
- Deterministic methods are inappropriate when the reservoir model contains sharp contrasts, e.g. binary facies models.
- The use of deterministic methods induces the risk of creating artifacts in the model, especially if a too loose regularization is applied.

A simple strategy where gradual deformation is used to pick out good initial guesses for the deterministic method is presented. This method can be used to minimize some of the drawbacks of the deterministic methods just mentioned. However, the methodology does not guarantee the geological consistency of the result.

In connection with the project the following programs have been developed:

- Boss reservoir simulator
  - The in-house reservoir simulator BOSS has been developed as a research framework for the IVC-SEP group, Technical University of Denmark. The presented history matching examples are all based on simulations with this simulator, except the cases related to streamline-derived regions.
- Adjoint extension to BOSS
  - In order to investigate the use of deterministic methods the BOSS simulator was extended with an adjoint sensitivity calculation feature. The current implementation can calculate sensitivities of injector pressures, watercuts, and oil rates with respect to grid block permeabilities and porosities. This implementation is together with the PPM implementation the backbone of the proposed hybrid method.
- PPM implementation with gradient guidance
  - The PPM method has been implemented in a framework which allows history matching with and without gradient guidance. The implementation calls the S-GeMS (Remy 2004) program to carry out sequential simulations.
- Implementation of gradual deformation
  - A gradual deformation history matching procedure has been implemented. Sequential simulation of permeability is done using GSLIB (Deutsch & Journel 1998).

Chapter 10

Future work

Several methods for history matching of oil production have been implemented and validated in this dissertation. The most promising method is the probability perturbation method (PPM) because of its flexibility with respect to the reservoir geology. Inclusion of qualitative gradient information is made possible by the methodology presented in Chapter 8.

This chapter seeks to present some relevant future work related to the presented history matching methods. Emphasis will be put on the PPM methodology since it seems to be the most versatile method and because it has a great potential in solving practical history matching problems. The cases considered in the dissertation have been represented by a binary geology - high permeable facies and low permeable facies. The PPM method works with any geology - binary, multiple categories as well as continuous variables. Inclusion of more than two facies is done by modifying the sequential simulation part of the history matching setup. This can be done by using a training image consisting of more than two categories. However, the use of the Snesim algorithm for simulation of multiple categorical reservoir models is RAM demanding. According to Journel (2006) the number of categories should not exceed four. With respect to the gradient-guided PPM method a number of issues are relevant to investigate:

- Applications to more realistic field cases
  - 3-phase systems
  - Larger models, i.e. more grid blocks
- Application to models with more than two categories

- Control strategy to indicate when to use gradient information
- Coupling to other sequential simulators, e.g. Filtersim, Simpat, etc.

The full potential of the gradient-guided PPM method will only be shown if the method is tested in a real-life history matching problem. It has been out of the scope of this work to include such a case. However, future work should be invested in more realistic cases. This could include three-phase systems and larger reservoir models. As already mentioned the work has been limited to deal with binary facies models. Two facies may be insufficient to describe a reservoir fully. Consequently, it is relevant to investigate the use of gradient guidance in cases with multiple facies types. This would necessitate a more advanced gradient scaling than the one used here. In the current setup the gradient is scaled according the the numerical value of the elements in the gradient. If the element is smaller than a specified negative threshold it is interpreted as an indication that the permeability should be increased at the particular grid block and vice versa if the value is larger than a positive threshold. For values in between the scaled gradient is set to zero. This scaling assumes that only two facies are present and an extension is needed if more than two facies types are included in the reservoir model. Inclusion of gradient information happens by a simple strategy in the current implementation. In every second or third outer iteration the gradient information is included. A more advanced strategy to control when to include the gradient information may be advantageous. Development of heuristic techniques may be a way to control the inclusion of gradient information better. The last point deals with the use of alternative sequential simulators. Snesim offers a fast and flexible way of constraining the match to a conceptual geological model. However, the pattern-based Simpat and Filtersim algorithms are known to reproduce complex geologic features better. The price is that the sequential simulation becomes more time demanding. Whether the increase in computation times compares with the computational load of performing a fluid flow simulation is questionable. Consequently, it is desirable to investigate the use of alternative sequential simulators in the proposed methodology.

With regards to the deterministic methods, future work could be invested in refining the regularization techniques used to stabilize the history matching problem. The deterministic methods offer an efficient means of calibrating a reservoir model to dynamic data. However, the efficiency is strongly dependent on the specific geology. A deterministic method is not suited for calibrating models with high contrasts, e.g. binary facies models.

The idea to use qualitative gradient information in history matching is an interesting subject. Inclusion of qualitative gradient knowledge into existing history matching work flows is an important research topic which could be the basis for future research.

Part III

Supplementary information

Appendix A

Details of adjoint approach

The mathematical background for the adjoint sensitivity calculation is given in Chapter 4, section 4.3. The purpose of this chapter is to describe the implementation of the adjoint methods in more detail.

A.1 Derivation of derivatives

The following sections describe the various derivatives used to calculate the sensitivities of production data with respect to grid block permeability and porosity. The production data covered are watercut and oil rate at producing wells and pressure of injection wells.

A.1.1 Derivatives with respect to simulation variables

This section deals with the derivation of the terms appearing in the adjoint equation (4.31). The derivatives needed in the adjoint equation are all with respect to the vector of unknowns, $\boldsymbol{u}$, i.e. with respect to grid block pressures and saturations. The term $(\nabla_{\boldsymbol{u}^n}(\boldsymbol{R}^n)^T)$ in equation (4.31) is the transposed Jacobian used in the solution of the non-linear discretized flow equations. Therefore, the derivation of this term is not included in this section since computation of the Jacobian as done in any reservoir simulator.

Derivation of $\frac{\partial \beta}{\partial \boldsymbol{u}}$, β = injector BHP

The total injection rate for an injection well can be found by summing equation (4.14) over the active segments:

$$q_{tot}^{i} = \sum_{j=1}^{N_{layers}} \lambda_{tot}^{i,j} WI_{i,j}(p_{blk}^{k} - p_{i,j}^{well}), \tag{A.1}$$

where k denotes the number of the grid block containing the jth segment of the ith injector.

Combination of equations (4.13) and (A.1) and reordering yields:

$$P_{inj}^{i} = \frac{1}{\sum_{j=1}^{N_{layers}} \lambda_{tot}^{k} WI_{i,j}} \cdot \left(q_{tot}^{i} - \sum_{j=1}^{N_{layers}} \lambda_{tot}^{k} WI_{i,j} p_k + g \sum_{j=2}^{N_{layers}} \left(\lambda_{tot}^{k} WI_{i,j} \sum_{m=2}^{j} \rho_w (z_m - z_{m-1}) \right) \right) \tag{A.2}$$

Note that k is the number of the grid block containing the jth segment and that k changes with j ($k = k(j)$).

In equation (A.2) it is assumed that the injected water can be treated as incompressible in the well. This is done in order to simplify the following derivations. With this assumption the derivative of injector bottomhole pressure with respect to pressure in a grid block containing a well segment is given as:

$$\frac{\partial P_{inj}^{i}}{\partial p_{k'}} = \frac{\lambda_{tot}^{k} WI_{i,j}}{\sum_{j=1}^{N_{layers}} \lambda_{tot}^{k} WI_{i,j}} \tag{A.3}$$

The derivative with respect to grid block saturations is:

$$\frac{\partial P_{inj}^{i}}{\partial S_{k'}} = \frac{1}{\left(\sum_{j=1}^{N_{layers}} \lambda_{tot}^{k} WI_{i,j} \right)^2} \left(\frac{\partial \lambda_{tot}^{k'}}{\partial S_{k'}} WI_{i,j} q_{tot}^{i} + \frac{\partial \lambda_{tot}^{k'}}{\partial S_{k'}} WI_{i,j} p_{k'} \sum_{j=1}^{N_{layers}} \lambda_{tot}^{k} WI_{i,j} - \frac{\partial \lambda_{tot}^{k'}}{\partial S_{k'}} WI_{i,j} \sum_{j=1}^{N_{layers}} \lambda_{tot}^{k} WI_{i,j} p_k + DOBSUM \right), \tag{A.4}$$

where $DOBSUM$ is given as:

$$DOBSUM = g\frac{\partial\lambda_{tot}^{k'}}{\partial S_{k'}}WI_{i,j}\sum_{m=2}^{j}\rho_w(z_m - z_{m-1})\sum_{j=1}^{N_{layers}}\lambda_{tot}^{k}WI_{i,j} - g\sum_{j=2}^{N_{layers}}\lambda_{tot}^{k}WI_{i,j}\sum_{m=2}^{j}\rho_w(z_m - z_{m-1})\frac{\partial\lambda_{tot}^{k'}}{\partial S_{k'}}WI_{i,j} \tag{A.5}$$

Derivation of $\frac{\partial\beta}{\partial \boldsymbol{u}}$, β = watercut

The watercut of a producer is given by the total water production divided with the total production (water and oil). The individual flow rates of water and oil into a well segment are given by equation (4.14). Watercut for the ith producer is given by:

$$WCUT_i = \frac{\sum_{j=1}^{N_{layers}}\frac{\rho_w}{\rho_w^{std}}\lambda_w WI_{i,j}(p_k^{blk} - p_{i,j}^{well})}{\sum_{j=1}^{N_{layers}}\frac{\rho_w}{\rho_w^{std}}\lambda_w WI_{i,j}(p_k^{blk} - p_{i,j}^{well}) + \sum_{j=1}^{N_{layers}}\frac{\rho_o}{\rho_o^{std}}\lambda_o WI_{i,j}(p_k^{blk} - p_{i,j}^{well})}, \tag{A.6}$$

or in compressed notation:

$$WCUT_i = \frac{w}{w+o} \tag{A.7}$$

The derivative of the water rate with respect to pressure in the kth grid block is given as:

$$\frac{\partial w}{\partial p_k} = \frac{WI_{i,j}}{\rho_w^{std}}(\frac{\partial\rho_w^k}{\partial p_k}\lambda_w^k + \frac{\partial\lambda_w^k}{\partial p_k}\rho_w^k)(p_k^{blk} - p_{i,j}^{well}) + \frac{WI_{i,j}}{\rho_w^{std}}\rho_w^k\lambda_w^k \tag{A.8}$$

And similarly for the oil rate:

$$\frac{\partial o}{\partial p_k} = \frac{WI_{i,j}}{\rho_o^{std}}(\frac{\partial\rho_o^k}{\partial p_k}\lambda_o^k + \frac{\partial\lambda_o^k}{\partial p_k}\rho_o^k)(p_k^{blk} - p_{i,j}^{well}) + \frac{WI_{i,j}}{\rho_o^{std}}\rho_o^k\lambda_o^k \tag{A.9}$$

The derivative of watercut with respect to pressure is then given as:

$$\frac{\partial WCUT_i}{\partial p_k} = \frac{\frac{\partial w}{\partial p_k}(w+o) - (\frac{\partial w}{\partial p_k} + \frac{\partial o}{\partial p_k})w}{(w+o)^2} \tag{A.10}$$

The derivative of water rate with respect to grid block saturation is given as:

$$\frac{\partial w}{\partial s_k} = \frac{\rho_w^k}{\rho_w^{std}}(p_k^{blk} - p_{i,j}^{well})WI_{i,j}\frac{\partial\lambda_w^k}{\partial S_k} \tag{A.11}$$

and the derivative of oil rate with respect to grid block saturation:

$$\frac{\partial o}{\partial s_k} = \frac{\rho_o^k}{\rho_o^{std}}(p_k^{blk} - p_{i,j}^{well})WI_{i,j}\frac{\partial \lambda_o^k}{\partial S_k} \tag{A.12}$$

The derivative of watercut with respect to grid block saturation is thus given as:

$$\frac{\partial WCUT_i}{\partial S_k} = \frac{\frac{\partial w}{\partial S_k}(w+o) - (\frac{\partial w}{\partial S_k} + \frac{\partial o}{\partial S_k})w}{(w+o)^2} \tag{A.13}$$

Derivation of $\frac{\partial \beta}{\partial \boldsymbol{u}}$, β = oil rate

The total oil rate is given as:

$$q_{tot,i}^o = \sum_{j=1}^{N_{layers}} \frac{\rho_o}{\rho_o^{std}}\lambda_o WI_{i,j}(p_k^{blk} - p_{i,j}^{well}) \tag{A.14}$$

The derivative of oil rate with respect to grid block pressure is:

$$\frac{\partial q_{tot,i}^o}{\partial p_k} = \frac{WI_{i,j}}{\rho_o^{std}}(\frac{\partial \rho_o^k}{\partial p_k}\lambda_o^k + \frac{\partial \lambda_o^k}{\partial p_k})(p_k^{blk} - p_{i,j}^{well}) + \frac{\rho_o^k \lambda_o^k WI_{i,j}}{\rho_o^{std}} \tag{A.15}$$

and the derivative with respect to grid block saturation is given as:

$$\frac{\partial q_{tot,i}^o}{\partial S_k} = \frac{WI_{i,j}\rho_o^k}{\rho_o^{std}}(p_k^{blk} - p_{i,j}^{well})\frac{\partial \lambda_o^k}{\partial S_k} \tag{A.16}$$

Derivation of $(\nabla_{\boldsymbol{u}^n}(\boldsymbol{R}^{n+1})^T)$ for equation (4.31)

In the implicit formulation of the discretized flow equations only the accumulation term involves properties evaluated in the current time step. All other terms in the discretized equations are evaluated implicitly in the in the next time step. The accumulation term is given as:

$$\boldsymbol{D}(\boldsymbol{u}^{n+1} - \boldsymbol{u}^n) = \frac{\boldsymbol{V}}{\Delta t}(\boldsymbol{\phi\rho S}|_{t+\Delta t} - \boldsymbol{\phi\rho S}|_t) \tag{A.17}$$

The derivative of R^{n+1} with respect to grid block pressure is:

$$\frac{\partial R_{2k-1}^{n+1}}{\partial p_k} = \frac{V_k}{\Delta t}\phi S_k \frac{\partial \rho_w^k}{\partial p_k} \tag{A.18a}$$

$$\frac{\partial R_{2k}^{n+1}}{\partial p_k} = \frac{V_k}{\Delta t}\phi(1 - S_k)\frac{\partial \rho_o^k}{\partial p_k} \tag{A.18b}$$

And the derivative with respect to grid block saturation is:

$$\frac{\partial R_{2k-1}^{n+1}}{\partial S_k} = \frac{V_k}{\Delta t}\phi\rho_w^k \tag{A.19a}$$

$$\frac{\partial R_{2k}^{n+1}}{\partial S_k} = -\frac{V_k}{\Delta t}\phi\rho_o^k \tag{A.19b}$$

A.1.2 Derivatives with respect to permeability

The sensitivity equation (4.33) contains derivatives with respect to the model parameter $\boldsymbol{m}$. In the following sections the derivatives needed to compute the sensitivity with respect to grid block permeabilities are supplied.

Derivation of $\frac{\partial\beta}{\partial\boldsymbol{m}}$, β = Injector BHP and $\boldsymbol{m}$ = permeability

In the sensitivity equation (4.33) the term $\nabla_{\boldsymbol{m}}\beta$ appears. This term accounts for explicit dependence of the production data with respect to the property $\boldsymbol{m}$. The injector bottomhole pressure as given in equation (A.2) is differentiated with respect to grid block permeability. Isotropic permeability is assumed, i.e. $k_x = k_y = k_z$.

$$\begin{aligned}\frac{\partial P_{inj}^i}{\partial k'} =& \frac{1}{\left(\sum_{j=1}^{N_{layers}} \lambda_{tot}^k WI_{i,j}\right)^2}\Bigg(q_{tot}^i\lambda_{tot}^{k'}\frac{\partial WI_{i,k}}{\partial k'} \\ &-\lambda_{tot}^{k'}p_{k'}\frac{\partial WI_{i,k}}{\partial k'}\sum_{j=1}^{N_{layers}}\lambda_{tot}^k WI_{i,j} + \lambda_{tot}^{k'}\frac{\partial WI_{i,k}}{\partial k'}\sum_{j=1}^{N_{layers}}\lambda_{tot}^k WI_{i,j}p_k - LT\Bigg),\end{aligned} \tag{A.20a}$$

where LT is given as:

$$\begin{aligned}LT =& g\lambda_{tot}^{k'}\frac{\partial WI_{i,k}}{\partial k'}\sum_{m=2}^{j'}\rho_w(z_m - z_{m-1})\sum_{j=1}^{N_{layers}}\lambda_{tot}^k WI_{i,j} - \\ &\lambda_{tot}^{k'}\frac{\partial WI_{i,k}}{\partial k'}\left(g\sum_{j=2}^{N_{layers}}\lambda_{tot}^k WI_{i,j}\left(\sum_{m=2}^{j}\rho_w(z_m - z_{m-1})\right)\right)\end{aligned} \tag{A.20b}$$

The notation j' in the first sum appearing in equation (A.20b) indicates a summation over overlying well segments.

Derivation of $\frac{\partial \beta}{\partial \boldsymbol{m}}$, β = watercut and $\boldsymbol{m}$ = permeability

The watercut is given as:

$$WCUT_i = \frac{\sum_{j=1}^{N_{layers}} \frac{\rho_w}{\rho_w^{std}} \lambda_w WI_{i,j}(p_k^{blk} - p_{i,j}^{well})}{\sum_{j=1}^{N_{layers}} \frac{\rho_w}{\rho_w^{std}} \lambda_w WI_{i,j}(p_k^{blk} - p_{i,j}^{well}) + \sum_{j=1}^{N_{layers}} \frac{\rho_o}{\rho_o^{std}} \lambda_o WI_{i,j}(p_k^{blk} - p_{i,j}^{well})}, \tag{A.6}$$

or in compressed notation:

$$WCUT_i = \frac{w}{w+o}$$

The derivative of water rate with respect to grid block permeability is:

$$\frac{\partial w}{\partial k} = \frac{\rho_w^k \lambda_w^k}{\rho_w^{std}}(p_k^{blk} - p_{i,j}^{well}) \frac{\partial WI_{i,k}}{\partial k} \tag{A.21}$$

and analogously for the oil rate

$$\frac{\partial o}{\partial k} = \frac{\rho_o^k \lambda_o^k}{\rho_o^{std}}(p_k^{blk} - p_{i,j}^{well}) \frac{\partial WI_{i,k}}{\partial k} \tag{A.22}$$

With these derivatives the derivative of watercut for the ith producing well is given as:

$$\frac{\partial WCUT_i}{\partial k} = \frac{\frac{\partial w}{\partial k}(w+o) - (\frac{\partial w}{\partial k} + \frac{\partial o}{\partial k})w}{(w+o)^2} \tag{A.23}$$

Derivation of $\frac{\partial \beta}{\partial \boldsymbol{m}}$, β = oil rate and $\boldsymbol{m}$ = permeability

The total oil rate for the ith producer is given as:

$$q_{tot,i}^o = \sum_{j=1}^{N_{layers}} \frac{\rho_o}{\rho_o^{std}} \lambda_o WI_{i,j}(p_k^{blk} - p_{i,j}^{well}) \tag{A.14}$$

The derivative with respect to grid block permeability is:

$$\frac{\partial q_{tot,i}^o}{\partial k} = \frac{\rho_o^k \lambda_o^k}{\rho_o^{std}}(p_k^{blk} - p_{i,j}^{well}) \frac{\partial WI_{i,j}}{\partial k} \tag{A.24}$$

Derivation of the $\nabla_{\boldsymbol{m}}(\boldsymbol{R}^n)$ term for $\boldsymbol{m}$ = permeability

The term $\nabla_{\boldsymbol{m}}(\boldsymbol{R}^n)$ in the sensitivity equation (4.33) plays an important role since it relates changes to the property $\boldsymbol{m}$ to changes in the residual. Permeability enters the discretized flow equations through the geometric transmissibility defined in equation (4.19b):

$$\Gamma_{i+1/2,j} = \frac{H_{res}\Delta y_j k_{x,i+1/2,j}}{\Delta x_{i+}}, \tag{4.19b}$$

with the harmonic average of the interface permeability given by equation (4.20a):

$$k_{x,i+1/2,j} = \frac{\Delta x_i + \Delta x_{i+1}}{\dfrac{\Delta x_i}{k_{x,i,j}} + \dfrac{\Delta x_{i+1}}{k_{x,i+1,j}}}. \tag{4.20a}$$

The derivative of the water flux term with respect to grid block permeability is:

$$\frac{\partial q^F_{i+1/2,j}}{\partial k_{x,i,j}} = H^w_{i+1/2,j}(\Delta x_i + \Delta x_{i+1})\frac{H_{ref}\Delta y_i}{\Delta x_+}\frac{k^2_{x,i+1,j}\Delta x_i}{(k_{x,i,i}\Delta x_{i+1} + k_{x,i+1,j}\Delta x_i)^2}(p_{i+1,j} - p_{i,j}) \tag{A.25a}$$

$$\frac{\partial q^F_{i+1/2,j}}{\partial k_{x,i+1,j}} = H^w_{i+1/2,j}(\Delta x_i + \Delta x_{i+1})\frac{H_{ref}\Delta y_i}{\Delta x_+}\frac{k^2_{x,i,j}\Delta x_{i+1}}{(k_{x,i,i}\Delta x_{i+1} + k_{x,i+1,j}\Delta x_i)^2}(p_{i+1,j} - p_{i,j}) \tag{A.25b}$$

A similar expression is obtained for the oil flux terms.

The number of the grid block with coordinates (i, j) is denoted n and the number of the grid block associated with the coordinates $(i+1, j)$ is denoted m. With this notation the derivative of the residual with respect to permeability in the two grid blocks is defined as:

$$\frac{\partial R_{2n-1}}{\partial k_{x,i,j}} = H^w_{i+1/2,j}(\Delta x_i + \Delta x_{i+1})\frac{H_{ref}\Delta y_i}{\Delta x_+}\frac{k^2_{x,i+1,j}\Delta x_i}{(k_{x,i,i}\Delta x_{i+1} + k_{x,i+1,j}\Delta x_i)^2}(p_{i+1,j} - p_{i,j}) \tag{A.26a}$$

$$\frac{\partial q^F_{2n-1}}{\partial k_{x,i+1,j}} = H^w_{i+1/2,j}(\Delta x_i + \Delta x_{i+1})\frac{H_{ref}\Delta y_i}{\Delta x_+}\frac{k^2_{x,i,j}\Delta x_{i+1}}{(k_{x,i,i}\Delta x_{i+1} + k_{x,i+1,j}\Delta x_i)^2}(p_{i+1,j} - p_{i,j}) \tag{A.26b}$$

$$\frac{\partial R_{2n}}{\partial k_{x,i,j}} = H^o_{i+1/2,j}(\Delta x_i + \Delta x_{i+1})\frac{H_{ref}\Delta y_i}{\Delta x_+}\frac{k^2_{x,i+1,j}\Delta x_i}{(k_{x,i,i}\Delta x_{i+1} + k_{x,i+1,j}\Delta x_i)^2}(p_{i+1,j} - p_{i,j}) \tag{A.26c}$$

$$\frac{\partial q^F_{2n}}{\partial k_{x,i+1,j}} = H^o_{i+1/2,j}(\Delta x_i + \Delta x_{i+1})\frac{H_{ref}\Delta y_i}{\Delta x_+}\frac{k^2_{x,i,j}\Delta x_{i+1}}{(k_{x,i,i}\Delta x_{i+1} + k_{x,i+1,j}\Delta x_i)^2}(p_{i+1,j} - p_{i,j}) \tag{A.26d}$$

Because of flux continuity over the grid block face the following expressions hold:

$$\frac{\partial R_{2m-1}}{\partial k_{x,i,j}} = -\frac{\partial R_{2n-1}}{\partial k_{x,i,j}} \tag{A.27a}$$

$$\frac{\partial q^F_{2m-1}}{\partial k_{x,i+1,j}} = -\frac{\partial q^F_{2n-1}}{\partial k_{x,i+1,j}} \tag{A.27b}$$

$$\frac{\partial R_{2m}}{\partial k_{x,i,j}} = -\frac{\partial R_{2n}}{\partial k_{x,i,j}} \tag{A.27c}$$

$$\frac{\partial q^F_{2m}}{\partial k_{x,i+1,j}} = -\frac{\partial q^F_{2n}}{\partial k_{x,i+1,j}} \tag{A.27d}$$

The last expressions come from the fact that the mass of the fluids leaving grid block n through the particular interface has to enter grid block m, i.e. $q_p^{F,n\to m} = -q_p^{F,m\to n}, p \in \{o, w\}$.

Sink/source terms

The sink and source terms need to be accounted for in the derivation of the $\nabla_{\boldsymbol{m}}(\boldsymbol{R}^n)$ term in equation (4.33).

The flow of water and oil into producer segment j under reservoir conditions can be computed by the following generalized expression:

$$q_w^{i,j} = -\rho_w^n \lambda_w^n W I_{i,j}(p_{blk}^n - p_{i,j}^{well})$$
$$q_o^{i,j} = -\rho_o^n \lambda_o^n W I_{i,j}(p_{blk}^n - p_{i,j}^{well}),$$

where n denotes the number of the grid block containing the j'th segment of the ith producer. p_{blk}^k is the grid block pressure of grid block k and $p_{i,j}^{well}$ is the well pressure in the segment computed by equation (4.13).

For production wells the derivative of the sink/source term entering into the residual is:

$$\frac{\partial R_{2n-1}}{\partial k_n} = \rho_w^n \lambda_w^n (p_{blk}^n - p_{i,j}^{well}) \frac{\partial W I_{i,j}}{\partial k_n} \tag{A.28a}$$

$$\frac{\partial R_{2n}}{\partial k_n} = \rho_o^n \lambda_o^n (p_{blk}^n - p_{i,j}^{well}) \frac{\partial W I_{i,j}}{\partial k_n} \tag{A.28b}$$

and for injections wells:

$$\frac{\partial R_{2n-1}}{\partial k_n} = \rho_w^n \lambda_{tot}^n (p_{blk}^n - p_{i,j}^{well}) \frac{\partial W I_{i,j}}{\partial k_n} \tag{A.29a}$$

$$\frac{\partial R_{2n}}{\partial k_n} = 0 \tag{A.29b}$$

It is assumed that only water is injected. If a mixture of water and oil is injected the term in equation (A.29a) has to be computed analogously to the water term.

Finally, the part in the residual related to the well equation has to be differentiated with respect to grid block permeability. p^k_{blk} is the grid block pressure of grid block k and $p^{well}_{i,j}$ is the well pressure in the segment computed by equation (4.13).

For the ith injection well the well equation can be generalized as:

$$\sum_{j=1}^{N_layers} \rho_w \lambda^n_{tot} WI_{i,j}(p^n_{blk} - p^{well}_{i,j}) - q^{i,tot}_w = 0, \tag{A.30}$$

where n denotes the number of the grid block containing the j'th segment of the ith producer.

The derivative of the well equation with respect to the permeability in the well segment becomes:

$$\frac{\partial R_{2N_{tot}+i}}{\partial k_n} = \rho^n_w \lambda^n_{tot}(p^n_{blk} - p^{well}_{i,j}) \frac{\partial WI_{i,j}}{\partial k_n}, \tag{A.31}$$

$$\tag{A.32}$$

where N_{tot} is the total number of grid blocks.

Derivative of well index

For a vertical well in a two-dimensional grid the well index for a well segment is given as:

$$WI = \left[\frac{\theta (k_x k_y)^{1/2} H}{\ln\left(\frac{r_o}{r_w}\right) + s} \right], \tag{4.12}$$

where k_x and k_y are the grid block permeabilities in the x-direction and y-direction, respectively. If isotropic permeability is assumed, i.e. $k_x = k_y = k$, the equation is simplified to:

$$WI = \left[\frac{\theta k H}{\ln\left(\frac{r_o}{r_w}\right) + s} \right], \tag{A.33}$$

and the derivative of the well index is simply:

$$\frac{\partial WI}{\partial k} = \left[\frac{\theta H}{\ln\left(\frac{r_o}{r_w}\right) + s} \right] = \frac{WI}{k}. \tag{A.34}$$

The definition $\frac{\partial WI}{\partial k} = \frac{WI}{k}$ is convenient when implementing the derivative in a reservoir simulator.

A.1.3 Derivatives with respect to porosity

Compared to the complexity and number of permeability derivatives discussed above the computation of the required porosity derivatives is fairly simple. Porosity does not enter into the computation of well pressures, watercuts, or production rates at the wells. Therefore, there is no explicit dependence in β in equation (4.33) on porosity. Consequently, the $\nabla_{\boldsymbol{m}}\beta$ terms vanish. Additionally, porosity does not enter into the computation of the inter grid block mass fluxes which simplifies the computation of the derivative of the residual with respect to porosity significantly. The only place in the discretized flow equation where porosity enters is in the computation of the accumulation term - through equation (4.16):

$$\boldsymbol{D}(\boldsymbol{u}^{n+1} - \boldsymbol{u}^n) = \frac{\boldsymbol{V}}{\Delta t}(\boldsymbol{\phi\rho S}|_{t+\Delta t} - \boldsymbol{\phi\rho S}|_t) \tag{4.16}$$

The derivative of the residual with respect to porosity in the nth grid block is then is then:

$$\frac{\partial R_{2n-1}}{\partial \phi_n} = \frac{V^n}{\Delta t}(\rho_w^n S^n|_{t+\Delta t} - \rho_w^n S^n|_t) \tag{A.35a}$$

$$\frac{\partial R_{2n}}{\partial \phi_n} = \frac{V^n}{\Delta t}(\rho_o^n (1-S^n)|_{t+\Delta t} - \rho_o^n (1-S^n)|_t) \tag{A.35b}$$

$$\frac{\partial R_{2n-1}}{\partial \phi_m} = 0, \text{for } n \neq m \tag{A.35c}$$

$$\frac{\partial R_{2n}}{\partial \phi_m} = 0, \text{for } n \neq m \tag{A.35d}$$

A.2 Implementational details

In the adjoint equation (4.31) the transposed Jacobian of the residual appears. In order to solve the adjoint equation the Jacobian has either to be stored for each time step in the reservoir simulation or information on the grid block saturations and pressures must be saved such that the Jacobian can be recomputed. In the adjoint implementation in BOSS the latter approach is used. Consequently, the vector of unknowns, $\boldsymbol{u}$, is stored for each time step in the simulation if an adjoint sensitivity calculation is wanted. A set of adjoint variables exist for each measured production data at each measurement time. The adjoint variables for the ith production data, at the jth measurement time will be denoted $\boldsymbol{\lambda}_{i,j}$. The index i denotes the type of production data, i.e. injector bottomhole pressure, watercut, or oil rate for each well.

The vector of adjoint variables is given by equation (4.31). The adjoint variables are determined by stepping backwards in time starting from the last time where a production data has been measured. In order to solve for the

adjoint variables for timestep n, $\boldsymbol{\lambda}^n$, we need the adjoint variables for the next time step, $\boldsymbol{\lambda}^{n+1}$ since these appear in the adjoint equation (4.31). Consequently, the present adjoint variables and the adjoint variables solved for in the last solve need to be stored simultaneously. A set of adjoint variables exist for each time step. However, it is not necessary to store them all in the strategy used in the BOSS implementation. When the set of adjoint variables, $\boldsymbol{\lambda}^n$ have been computed they can be used to compute one of the terms in the sum appearing in the sensitivity equation (4.33). The general workflow for the computation of sensitivities is sketched out below. The number of measured production data is denoted N_p. If watercut and oil rate is measured at N_{prod} producers and injection bottomhole pressure is measured at N_{inj} injectors, N_p will be $2N_{prod}+N_{inj}$. The number of times where the production data is measured is denoted N_{mes}. When a non-linear-least squares method is used for history matching a number of N_pN_{mes} sensitivities are required to form the an approximation to the Hessian - see Chapter 7 for the details. As mentioned, only two sets of adjoint variables need to be stored. In the following sketch-out of the workflow the number of adjoint variables is therefore limited to two sets: $\boldsymbol{\lambda}_{i,j}^{old}$ and $\boldsymbol{\lambda}_{i,j}$, where $\boldsymbol{\lambda}_{i,j}^{old}$ denotes the adjoint variables solved for in the previous solve and $\boldsymbol{\lambda}_{i,j}$ denotes the new set.

1. Perform a reservoir simulation. Save grid block pressures and saturations for each time step.
2. Initiate $\lambda_{i,j}^{old} = \mathbf{0}, \quad i \in [1, N_p], j \in [1, N_{mes}]$.
3. Set t equal to the time of the last measurement and set n equal to the corresponding number of time steps.
4. Read the stored pressures and saturations for the time t and compute the Jacobian.
5. If t coincides with a measurement time compute $\nabla_{\boldsymbol{u}^n}\beta_{i,j}$ for all i and set $N_{eq} = N_{eq} + N_p$ and $N_t = N_t + 1$.
6. Solve the N_{eq} adjoint equations with equation (4.31) and get $\boldsymbol{\lambda}_{i,j} \quad i \in [1, N_p], j \in [1, N_t]$.
7. If t coincides with a measurement time compute $\nabla_{\boldsymbol{m}}\beta_{i,j}$.
8. Compute the N_{eq} $(\boldsymbol{\lambda})^T(\nabla_{\boldsymbol{m}}(\boldsymbol{R}^n)^T)$ terms in equation (4.33) and add $\nabla_{\boldsymbol{m}}\beta_{i,j}$.
9. Add the sensitivity terms to the sensitivity vectors: $s_{i,j} = s_{i,j} + (\boldsymbol{\lambda}_{i,j})^T(\nabla_{\boldsymbol{m}}(\boldsymbol{R}^n)^T) \quad i \in [1, N_p], j \in [1, N_t]$.
10. Set $\boldsymbol{\lambda}_{i,j}^{old} = \boldsymbol{\lambda}_{i,j}$

11. $n = n - 1$
12. $t = t - \Delta t$
13. If $n > 1$ goto 4, else END

If a direct solver has been chosen for the solution of the adjoint equations a factorization of the Jacobian is performed in each step. This is by far the most expensive operation in the method. Hence, the backwards solve is comparable to a reservoir simulation with respect the computation time where the factorization is also the most expensive operation. An obvious improvement would be to store the factorized Jacobian every time the Jacobian is refactorized in the forward simulation. The non-linear solver in the forward simulation reuses the factorized Jacobian until the convergence deteriorates. A refactorization of the Jacobian is then performed. Consequently, it is likely that the factorized Jacobian cannot be stored for each time step. However, the factorized Jacobian is needed for each time step for the solution of the (linear) adjoint equation. A possible solution may be to apply the so-called *staggered corrector* method suggested by Feehery, Tolsma & Barton (1997) who suggest the use of a quasi-Newton type method to solve a linear problem, where the system matrix is infrequently factorized. This has not been implemented in the current code since time did not allow for such extensions.

Appendix B

Dekker-Brent derivative-free optimizer

The Dekker-Brent optimization scheme is a derivative free optimization algorithm utilizing a hybrid golden section search and interpolation method. The method is robust and Brent (1973) guarantees it's convergence. The implementation is only made for one-dimensional optimizations. An excellent description of the algorithm is found in Brent (1973) and only the most important features are discussed here.

During the optimization six points are stored, these will be denoted a, b, u, v, w, x. The points a and b define the search interval in which the optimum is searched for. x is the point where the lowest function value has been calculated, w is the point representing the next lowest value, and v is the previous value of w - finally, u is the point where the objective function was evaluated most recently. Some of the points may coincide. Figure B.1 shows how the six points may be distributed.

When the optimization is initiated some of the points are initiated as

$$v = w = x = a\left(\frac{3-\sqrt{5}}{2}\right)(b-a). \tag{B.1}$$

The term $\frac{3-\sqrt{5}}{2} \approx 0.382$ ensures that the first step corresponds to a golden section step. Another factor may be used if it is deemed more reasonable. In the present implementation, however, only golden section search is used.

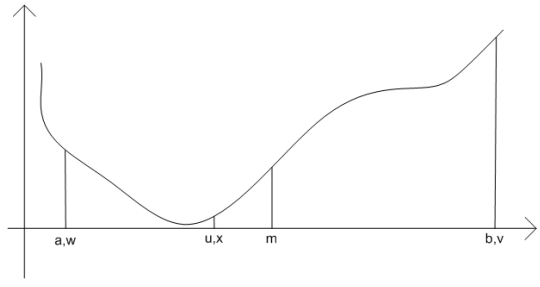


Figure B.1: A possible distribution of the six points and the midpoint indicated with m.

An important parameter in the algorithm is the tolerance given as

$$tol = eps|x| + t, \tag{B.2}$$

where first term is a relative tolerance and the last term is the absolute tolerance. The routine will return the minimizer to an accuracy of less then $3tol$. Let the mid point of the interval be $\frac{1}{2}(a+b)$. If $|x-m| \leq 2tol - \frac{1}{2}(b-a)$, or, equivalently, $max(x-a, b-x) \leq 2tol$, the procedure terminates and returns x as the minimizer. If the procedure does not terminate two numbers are calculated:

$$p = \pm[(x-v)^2(f(x)-f(w)) - (x-w)^2(f(x)-f(v))] \tag{B.3}$$
$$q = \mp 2[(x-v)(f(x)-f(w)) - (x-w)(f(x)-f(v))] \tag{B.4}$$

The numbers p and q are defined such that the point $x + \frac{p}{x}$ is the location of the turning point of the parabola passing through the points $(v, f(v))$, $(w, f(w))$, and $(x, f(x))$. The point $x + \frac{p}{q}$ is thus used as the new value of the independent variable (i.e. the point u). All this corresponds to a parabolic interpolation step.

Let e be the value of $\frac{p}{q}$ from the second-last cycle. If $|e| \leq tol$ or q=0 or $x + \frac{p}{q} \notin [a; b]$ or $|\frac{p}{q}| \geq \frac{1}{2}|e|$ the polynomial step is not accepted and a golden section step is taken instead, as

$$u = \begin{cases} \left(\frac{\sqrt{5}-1}{2}\right) x + \left(\frac{3-\sqrt{5}}{2}\right) a & \text{if } x \geq m, \\ \left(\frac{\sqrt{5}-1}{2}\right) x + \left(\frac{3-\sqrt{5}}{2}\right) b & \text{if } x < m. \end{cases} \tag{B.5}$$

The original strategy presented by Brent (1973) is used in the implementations of gradual deformation and the probability perturbation method. Only one thing is corrected: After initialization one point is always known. If the perturbation parameter or deformation parameter is equal to zero a reservoir

simulation will result in the current best value of the objective function. Therefore, the value of x is initialized to zero and not to the golden section step as stated in equation B.1.

Appendix C

The Levenberg-Marquardt method

The Levenberg-Marquardt (LM) method is used in Chapter 7 where it is applied to history matching problems. This appendix provides some background information on the LM method together with some implementational issues for large scale problems. The derivation of the Levenberg-Marquardt equations is based on the derivations given in Madsen, Nielsen & Tingleff (1999).

C.1 Gauss-Newton and Levenberg-Marquardt

We will consider the following generic optimization problem

$$\boldsymbol{x}^* = \underset{\boldsymbol{x}}{\text{Argmin}} \left[F(\boldsymbol{x}) \right], \tag{C.1}$$

where the objective function, F, is defined as a least-squares measure of the residual vector denoted $\boldsymbol{f}$:

$$F(\boldsymbol{x}) = \frac{1}{2} \boldsymbol{f}(\boldsymbol{x})^T \boldsymbol{f}(\boldsymbol{x}). \tag{C.2}$$

In a history matching context the residual contains the production data mismatches.

A Taylor expansion of the residual yields

$$\boldsymbol{f}(\boldsymbol{x} + \boldsymbol{h}) = \boldsymbol{f}(\boldsymbol{x}) + \boldsymbol{J}(\boldsymbol{x})\boldsymbol{h} + \mathcal{O}(\|\boldsymbol{h}\|^2), \tag{C.3}$$

where $\boldsymbol{J}$ it the Jacobian matrix defined as:

$$\boldsymbol{J}(\boldsymbol{x})_{i,j} = \frac{\partial f_i}{\partial x_j} \tag{C.4}$$

For small $\|\boldsymbol{h}\|$ the following linear model of $\boldsymbol{f}$ is valid:

$$\boldsymbol{f}(\boldsymbol{x}+\boldsymbol{h}) \approx \boldsymbol{l}(\boldsymbol{x}) \equiv \boldsymbol{f}(\boldsymbol{x}) + \boldsymbol{J}(\boldsymbol{x})\boldsymbol{h} \tag{C.5}$$

The linearized model for $\boldsymbol{f}$ is inserted into equation (C.2) which yields a linearized model for the objective function:

$$F(\boldsymbol{x}+\boldsymbol{h}) \approx L(\boldsymbol{h}) \equiv F(\boldsymbol{x}) + \boldsymbol{h}^T\boldsymbol{J}^T\boldsymbol{f} + \frac{1}{2}\boldsymbol{h}^T\boldsymbol{J}^T\boldsymbol{J}\boldsymbol{h} \tag{C.6}$$

L is at a minimum when it's derivative with respect to $\boldsymbol{h}$ is 0, i.e. when

$$\boldsymbol{J}^T\boldsymbol{f} + \boldsymbol{J}^T\boldsymbol{J}\boldsymbol{h} = 0. \tag{C.7}$$

The step which minimizes L is denoted the Gauss-Newton step and is given as

$$-\boldsymbol{J}^T\boldsymbol{f} = \boldsymbol{J}^T\boldsymbol{J}\boldsymbol{h}_{GN}. \tag{C.8}$$

It can be shown that $\boldsymbol{h}_{GN}$ is a descent direction. In the Gauss-Newton method the vector of independent variables is updated with the Gauss-Newton step, i.e. $\boldsymbol{x}^{n+1} = \boldsymbol{x}^n + \boldsymbol{h}_{GN}$. Coupled with a line search the method will have guaranteed convergence if the Jacobian has full rank. If the step becomes too large the linearization in (C.6) will no longer be valid and convergence will deteriorate.

The Levenberg-Marquardt method is a modified Gauss-Newton method where the update is given as

$$(\boldsymbol{J}^T\boldsymbol{J}\mu\boldsymbol{I})\boldsymbol{h}_M = -\boldsymbol{J}^T\boldsymbol{f}, \tag{C.9}$$

where $\boldsymbol{I}$ denotes the identity matrix and μ is a scalar referred to as the damping parameter. The following observations are central:

- If μ is large the Levenberg-Marquardt step is approximately given as $\boldsymbol{h}_M \approx -\frac{1}{\mu}\boldsymbol{J}\boldsymbol{f}$, i.e. a small step in the steepest descent direction.
- If μ is small the Levenberg-Marquardt step is approximate equal to the Gauss-Newton step, i.e. $\boldsymbol{h}_M \approx \boldsymbol{h}_{GN}$

The Levenberg-Marquardt method does not need a line search because the step length is controlled by the damping parameter. The value of the damping parameter is updated as the optimization progresses. In the implementation used for this dissertation the empirical update of the damping parameter described in Madsen et al. (1999) is used.

C.1.1 The Levenberg-Marquardt equation in the sparse formulation

The key step in the algorithm is the computation of the Levenberg-Marquardt (LM) step, denoted as $\boldsymbol{h}_M$:

$$(\boldsymbol{A} + \mu \boldsymbol{I})\boldsymbol{h}_M = -\boldsymbol{g}, \tag{C.10}$$

where $\boldsymbol{A}$ is the Gauss-Newton approximation to the Hessian, i.e. $\boldsymbol{A} = \boldsymbol{J}^T\boldsymbol{J}$ and $\boldsymbol{g}$ is the gradient given as $\boldsymbol{J}^T\boldsymbol{f}$ with $\boldsymbol{f}$ being the residual. In regularized problems the Jacobian consists of two parts, one part is from the sensitivities of the production data, the other is related to the regularization. The Jacobian can be represented as:

$$\boldsymbol{J} = \begin{bmatrix} \boldsymbol{J}_p \\ \boldsymbol{J}_r \end{bmatrix} \tag{C.11}$$

The upper part of $\boldsymbol{J}$ is a full matrix but the part related to regularization is sparse since the only nonzero elements in $\boldsymbol{J}_r$ are the diagonal elements. Therefore, it seems inappropriate to from the Hessian explicitly since this matrix will be a full matrix of size $N \times N$, N being the number of grid blocks. Equation (C.10) can be formulated as

$$\begin{bmatrix} \boldsymbol{f}_p(x) \\ \boldsymbol{r}(x) \\ \boldsymbol{0} \end{bmatrix} + \begin{bmatrix} \boldsymbol{J}_p(x) \\ \boldsymbol{J}_r(x) \\ \sqrt{\mu}\boldsymbol{I} \end{bmatrix} \boldsymbol{h}_m \approx \boldsymbol{0} \tag{C.12}$$

where $\boldsymbol{f}_p(x)$ denotes the production data mismatches and $\boldsymbol{r}(x)$ denotes the regularization mismatch.

The solution of this system is efficiently obtained by the use of the LSQR (Paige & Saunders 1982) solver which is specially designed to solve least squares problems.

List of Symbols

C	Covariance	
$\boldsymbol{C}$	Covariance matrix	
E	Objective function	N/D
$E(\cdot)$	Expected value	
$\boldsymbol{G}_{scl}$	Scaled gradient	Varying
$\boldsymbol{J}$	Jacobian	Varying
P	Pressure	Pa, $bars$
$P(A)$	Marginal probability of event A	N/D
$P(A\|B)$	Conditional probability of event A	N/D
Q	Source/Sink , see equation (4.2), page 31	kg/day
S_j	Saturation of phase j	N/D
Z	Depth	m
$\boldsymbol{Z}$	Realization of geological model	N/D
a	Degree of trust in the gradient	N/D
d	Observable production data	
g	Gravitational acceleration	m/s^2
$\boldsymbol{g}$	Gradient	Varying
$\boldsymbol{i}$	Binary facies indicator	N/D

$\mathbf{k}$	Permeability tensor	mDa
k_i	Absolute permeability in the i-direction	mDa
k_{rj}	Relative permeability of phase j	N/D
$\boldsymbol{r}$	Residual	Varying
r_c	Perturbation parameter	N/D
$\boldsymbol{s}$	Sensitivity vector	Varying
α_i	Weight of ith realization in gradual deformation	Varying
β	Production data in adjoint equations	
γ	Semivariogram	
λ	Relative phase mobility	$\frac{1}{Pa \cdot s}$
$\boldsymbol{\lambda}$	Adjoint variable	Varying
μ	Mean value	
μ	Viscosity	$Pa \cdot s$
ν	Anisotropi factor	
ρ	Deformation parameter	N/D
ρ	Density	kg/m^3
σ	Kriging variance	N/D
σ_r	Regularization parameter	N/D

List of Figures

List of Tables

Bibliography

Agarwal, B. & Blunt, M. J. (2003), 'Streamline-based method with full physics forward simulation for history-matching performance data of a north sea field', *Society of Peroleum Engineers* (84952).

Aziz, K., Durlofsky, L. & Tchelepi, H. (2005), 'Notes on petroleum reservoir simulation', Stanford University, California, USA.

Aziz, K. & Settari, A. (1979), *Petroleum Reservoir Simulation*, Elsevier Applied Science Publishers.

Barker, J. & Fayers, F. (1994), 'Transport coefficients for compositional simulation with coarse grids in heterogeneous media', *SPE Advanced Technology Series* **2**(2). SPE number 22591-PA.

Batycky, R. P. (1997), A Three-Dimensional Two-Phase Field Scale Streamline Simulator, PhD thesis, Stanford University.

Brent, R. P. (1973), *Algorithms for Minimization without Derivatives*, Series in Automatic Computation, Printice Hall.

Caers, J. (2003), 'Geostatistical history matching under training-image based geological constraints', *SPE Journal* **8**(3), 218–226.

Caers, J. (2007), 'Comparing the gradual deformation with the probability perturbation method for solving inverse problems', *Mathematical Geology* **39**(1).

Cheng, H., Kharghoria, A., He, Z. & Datta-Gupta, A. (2004), 'Fast history matching of finite-difference models using streamline-derived sensitivities', *Society of Petroleum Engineers* (89447).

Conradsen, K. (2003), *En Introduktion til Statistik*, Vol. 2, 6 edn, Informatics and Mathematical Modelling, Technical University of Denmark.

da Rocha, M. M. & Yamamoto, J. K. (2000), 'Comparison between kriging variance and interpolation variance as uncertainty measurements in the capanema iron mine, state of minas geraisbrazil', *Natural Resources Research* **9**(3), 223–235.

de Marsily, G., Lavedan, G., Boucher, M. & Fasanino, G. (1984), Interpretation of interference tests in a well field using geostatistical techniques to fit the permeability distribution in a reservoir model, *in* G. Verly, M. David & A. J. and. A. Marechal, eds, 'Geostatistics for natural resources characterization', Vol. 2, D. Reidel Publishing Company, pp. 831–849.

Deutsch, C. V. & Journel, A. G. (1998), *GSLIB - Geostatistical Software Library and User's Guide*, 2 edn, Oxford University Press.

Feehery, W. E., Tolsma, J. E. & Barton, P. I. (1997), 'Efficient sensitivity analysis of large-scale differential-algebraic systems', *Applied Numerical Mathematics* (25), 41–54.

Goovaerts, P. (1997), *Geostatistics for Natural Resources Evaluation*, Applied Geostatistics Series, Oxford University Press.

Haber, E., Ascher, U. M. & Oldenburg, D. (2000), 'On optimization techniques for solving nonlinear inverse problems', *Inverse Problems* (16), 1263–1280.

Hoffman, B. T. & Caers, J. (2003), 'Geostatistical history matching using a regional probability perturbation method', *Society of Petroleum Engineers* (SPE no. 84409).

Hohn, M. E. (1999), *Geostatistics and Petroleum Geology*, Kluwer Adademic Publishers.

Hu, L. Y. (2000), 'Gradual deformation and iterative calibration of gaussian-related stochastic models', *Mathematical Geology* **32**(1), 87–108.

Hu, L. Y. (2002), 'Combination of dependent realizations within the gradual deformation method', *Mathematical Geology* **34**(8), 953–963.

Hu, L. Y. & Jenni, S. (2005), 'History matching of object-based stochastic reservoir models', *SPE Journal* .

Hu, L. Y. & le Ravalec-Dupin, M. (2004), 'An improved gradual deformation method for reconciling random and gradient searches in stochastic optimizations', *Mathematical Geology* **36**(6), 703–719.

Isaaks, E. H. & Srivastava, R. M. (1989), *Applied Geostatistics*, Oxford University Press.

Johansen, K., Caers, J. & Suzuki, S. (2007), 'Hybridization of the probability perturbation method with gradient information', *Journal of Computational Geosciences* . To appear.

Johansen, K. & Kristensen, M. R. (2006), 'A two-dimensional two-phase black-oil reservoir simulator - the development of boss black-oil system simulator', Technical Report, IVC-SEP, Technical Universiy of Denmark.

Johansen, K., Shapiro, A. & Stenby, E. (2006), 'Statistical methods for history matching'. Prepared for the 26th conference on enhanced oil recovery organized by the International Energy Agency held in Paris, France.

Journel, A. G. (2002), 'Combining knowledge from diverse information sources: an alternative to bayesian analysis', *Mathematical Geology* **34**(5).

Journel, A. G. (2006), 'Lesson 12: An introduction to mp geostatistics'. Lecture notes from course PE/GES240, Department of Energy Resources Engineering, Stanford University.

Li, R., Reynolds, A. & Oliver, D. (2003), 'History matching of three-phase flow production data', *SPE Journal* (87336), 328–340.

Madsen, H. (1998), *Tidsrækkeanalyse*, 3 edn. Informatics and Mathematical Modelling, Technical University of Denmark.

Madsen, K., Nielsen, H. B. & Tingleff, O. (1999), 'Methods for non-linear least squares problems', Teaching material for course no. 02611 - Optimization and Data Fitting, Department of Mathematical Modelling, Technical University of Denmark.

Nocedal, J. & Wright, S. J. (1999), *Numerical Optimization*, Springer Series in Operations Research, Springer.

Nævdal, G., Johnsen, L. M., Aanonsen, S. I. & Vefring, E. H. (2003), 'Reservoir monitoring and continuous model updating using ensemble kalman filter', *SPE no. 84372* .

Nævdal, G., Mannseth, T. & Vefring, E. H. (2002), 'Near-well reservoir monitoring through ensemble kalman filter', *SPE no. 75235* .

P. Sarma, K. Aziz, L. J. D. (2005), 'Implementation of adjoint solution for optimal control of smart wells', *SPE no. 92864* .

Paige, C. C. & Saunders, M. A. (1982), 'Algorithm 583: Lsqr: Sparse linear equations and least squares problems', *ACM Trans. Math. Softw.* **8**(2).

Peaceman, D. W. (1983), 'Interpretation of well-block pressures in numerical reservoir simulation with non-square grid blocks and anisotropic permeability', *SPE Journal* **275**(1), 531–543.

Petruccelli, J., Nandram, B. & M.Chen (1999), *Applied Statistics for Engineers and Scientists*, Prentice Hall.

RamaRao, B. S., LaVenue, A. M., de Marsily, G. & Marietta, M. G. (1995), 'Pilot point methodology for automated calibration of an ensemble of conditionally simulated transmissivity fields', *Water Resources Research* **31**(3), 475–493.

Ravalec-Dupin, M. L., Nætinger, B., Hu, L.-Y. & Blanc, G. (2001), 'Condtioning to dynamic data: an improved zonation approach', *Petroleum Geoscience* **7**, S9–S16.

Remy, N. (2004), *Geostatistical Earth Modeling Software: Users Manual*, Stanford Center for Reservoir Forecasting (SCRF), Stanford University, California, USA.

Strebelle, S. (2000), Sequential simulation drawing structure from training images, PhD thesis, Stanford University, Stanford, California.

Strebelle, S. (2002), 'Conditional simulation of complex geological structures using multiple-point statistics', *Mathematical Geology* **34**(1).

Suzuki, S. & Caers, J. (2006), 'History matching with an uncertain geological scenario', *SPE 102154* .

Tikhonov, A. N. & Arsenin, V. Y. (1977), *Solutions of Ill-Posed Problems*, V. H. Winston & Sons, Washington, D.C.: John Wiley & Sons, New York. Translated from Russian, Preface by translation editor Fritz John, Scripta Series in Mathematics.

Wang, Y. & Kovcek, A. R. (2000), 'A streamline approach for history-matching production data', *Society of Petroleum Engineers* (59370).

Wu, Z., Reynolds, A. & Oliver, D. (1998), 'Conditioning geostatistical models to two-phase production data', *Society of Petroleum Engineers* (SPE No. 49003), 205–216.

Zhang, T., Bombarde, S., Strebelle, S. & Oatney, E. (2006), '3d porosity modeling of a carbonate reservoir using continuous multiple-point statistics simulation', *SPE Journal* **11**(3).

Department of Chemical
and Biochemical Engineering
DTU Building 229
Søltofts Plads
DK-2800 Kgs. Lyngby
www.kt.dtu.dk